

COK-A01210-27-P007688

American Journal of Agricultural Economics

Volume 73

Number 1

February 1991

Farmland Prices	Barry Falk
Campaign Contributions and Voting	David G. Abler
Aggregate Risk Effects	Jack Meyer & Lindon J. Robison
Computerized Instruction	David L. Debertin & Larry D. Jones
Taxes and Supply Response	John Quiggin
U.S. Beekeeping Industry	Lois Schertz Willett & Ben C. French
Grain Futures Options	William W. Wilson & Hung-Gay Fung
Options and Futures Markets ...	Harvey Lapan, Giancarlo Moschini & Steven D. Hanson
Dynamic Marketing Strategies	Russell Tronstad & C. Robert Taylor
Milk Supply Elasticities	Adesoji O. Adelaja
Spatial Trade Models	Giovanni Anania & Alex F. McCalla
Japanese Livestock Industry	Thomas I. Wahl, Dermot J. Hayes & Gary W. Williams
Competitiveness of Exports	Daniel H. Pick & Timothy A. Park
Export Subsidy Comparisons	Robert G. Chambers & Philip L. Paarlberg
Regulated Fisheries	Diane P. Dupont
Spatial Water Allocation	Ujjayant Chakravorty & James Roumasset
Control of Fish Growth	Oscar J. Cacho, Henry Kinnucan & Upton Hatch
Conservation Compliance	Dana L. Hoag & Herb A. Holloway
Rural Road Services	Steven C. Deller & Carl H. Nelson
Food Stamp Program	Barbara Devaney & Robert Moffitt
Demand for Fats and Oils	Brian W. Gould, Thomas L. Cox & Federico Perali



AAEA Annual Meeting
Kansas State University
Manhattan, Kansas
4-7 August 1991

Editorial Staff

Editor: Peter J. Barry, University of Illinois

Senior Associate Editors: John Braden and Philip Garcia, University of Illinois

Associate Editors: Philip C. Abbott, Purdue University; Julian M. Alston, University of California, Davis; Jock R. Anderson, University of New England; F. S. Bagi, Tennessee State University; Bruce R. Beattie, Montana State University; Steven T. Buccola, Oregon State University; Jean-Paul Chavas, University of Wisconsin; Bruce L. Dixon, University of Arkansas; B. Delworth Gardner, Brigham Young University; Peter B. R. Hazell, World Bank; Robert W. Herdt, Rockefeller Foundation; Thomas W. Hertel, Purdue University; Robert P. King, University of Minnesota; Linda K. Lee, University of Connecticut; William E. Martin, University of Illinois; Bruce A. McCarl, Texas A&M University; William H. Meyers, Iowa State University; Jeffrey M. Perloff, University of California, Berkeley; Daniel A. Sumner, North Carolina State University; Stanley R. Thompson, Michigan State University

Book Review Editor: Dale W. Adams, Ohio State University

Technical Editor: Martha A. Luzader, P.O. Box 189, Rockville IN 47872-0189. Phone (317) 344-1426

Editorial Assistant: Phyllis Blackford

Editorial Communications

All manuscripts and editorial correspondence, except as noted below, should be addressed to:

Peter J. Barry, Editor
Department of Agricultural Economics
University of Illinois
305 Mumford Hall
1301 W. Gregory Drive
Urbana, Illinois 61801

Effective 1 February 1991, submissions of new manuscripts should be sent to: Steven T. Buccola and Richard M. Adams, Co-editors, Department of Agricultural and Resource Economics, Oregon State University, Corvallis, OR 97330.

Communications concerning book reviews and books submitted for announcement in the *Journal* should be sent to the book review editor: Otto Doering, Department of Agricultural Economics, Krannert Building, Purdue University, West Lafayette, IN 47907.

Editorial Policy

The purpose of the *Journal* is to provide a forum for creative and scholarly work in agricultural economics. Thus, acceptable manuscripts should have a relationship to the economics of agriculture, natural resources, or rural and community development. Contributions, methodological or applied, in the business, extension, research, and teaching phases of agricultural economics are equally encouraged. The *American Journal of Agricultural Economics* (ISSN 0002-9092) is published monthly in February, May, August, November, and December by the American Agricultural Economics Association, 80 Heady Hall, Iowa State University, Ames, Iowa and printed by Edwards Brothers, Inc., Ann Arbor, MI. Second-class postage is paid at Ames, Iowa, and additional mailing offices. Postmaster: Send address changes to the American Journal of Agricultural Economics, 80 Heady Hall, Iowa State University, Ames, Iowa, 50011-1070.

Subscription

Subscription cost is \$65 per year plus foreign postage; members receive copies as part of their membership fee of \$60 per year.

Copyright 1991 American Agricultural Economics Association. Any article or other material appearing in the *American Journal of Agricultural Economics* may not be republished in full or in part without the written permission of the editor.

American Journal of Agricultural Economics

Volume 73

Number 1

February 1991

Articles

338.105
AM 35

Formally Testing the Present Value Model of Farmland Prices/*Barry Falk*/1

Campaign Contributions and House Voting on Sugar and Dairy Legislation/*David G. Abler*/11

The Aggregate Effects of Risk in the Agricultural Sector/*Jack Meyer and Lindon J. Robison*/18

Applications of Computer Graphics to Undergraduate Instruction in Agricultural Economics/*David L. Debertin and Larry D. Jones*/25

Supply Response Under Proportional Profits Taxation/*John Quiggin*/36

An Econometric Model of the U.S. Beekeeping Industry/*Lois Schertz Willett and Ben C. French*/40

Put-Call Parity and Arbitrage Bounds for Options on Grain Futures/*William W. Wilson and Hung-Gay Fung*/55

Production, Hedging, and Speculative Decisions with Options and Futures Markets/*Harvey Lapan, Giancarlo Moschini, and Steven D. Hanson*/66

Dynamically Optimal After-Tax Grain Storage, Cash Grain Sale, and Hedging Strategies/*Russell Tronstad and C. Robert Taylor*/75

Price Changes, Supply Elasticities, Industry Organization, and Dairy Output Distribution/*Adesoji O. Adelaja*/89

Does Arbitraging Matter? Spatial Trade Models and Discriminatory Trade Policies/*Giovanni Anania and Alex F. McCalla*/103

Dynamic Adjustment in the Japanese Livestock Industry Under Beef Import Liberalization/*Thomas I. Wahl, Dermot J. Hayes, and Gary W. Williams*/118

The Competitive Structure of U.S. Agricultural Exports/*Daniel H. Pick and Timothy A. Park*/133

Are More Exports Always Better? Comparing Cash and In-Kind Export Subsidies/*Robert G. Chambers and Philip L. Paarlberg*/142

Testing for Input Substitution in a Regulated Fishery/*Diane P. Dupont*/155

Efficient Spatial Allocation of Irrigation/*Ujjayant Chakravorty and James Roumasset*/165

Optimal Control of Fish Growth/*Oscar J. Cacho, Henry Kinnucan, and Upton Hatch*/174

Farm Production Decisions Under Cross- and Conservation Compliance/
Dana L. Hoag and Herb A. Holloway/184

Measuring the Economic Efficiency of Producing Rural Road Services/
Steven C. Deller and Carl H. Nelson/194

Dietary Effects of the Food Stamp Program/*Barbara Devaney and Robert Moffitt/*
202

Demand for Food Fats and Oils: The Role of Demographic Variables and
Government Donations/*Brian W. Gould, Thomas L. Cox, and Federico Perali/*
212

Books Reviewed

Bowes, Michael D., and John V. Krutilla. *Multiple-Use Management: The
Economics of Public Forestlands/Richard J. Brazee/222*

Braden, John B., and Stephen B. Lovejoy, eds. *Agriculture and Water Quality:
International Perspective/Clyde F. Kiker/223*

Cohn, Theodore H. *The International Politics of Agricultural Trade: Canadian-
American Relations in a Global Agricultural Context/Leo V. Mayer/224*

Fuguitt, Glenn V., David L. Brown, and Calvin L. Beale. *Rural and Small
Town America/C. Shannon Stokes/225*

James, W. E., S. Naya, and G. M. Meier. *Asian Development: Economic
Success and Policy Lessons/Alfred Thieme, Jr./226*

Moore, Richard H. *Japanese Agriculture, Patterns of Rural Development/
Fred H. Sanderson/227*

Pearce, D. W., and R. K. Turner. *Economics of Natural Resources and the
Environment/Jeff W. Bennett/227*

Rawski, Thomas G. *Economic Growth in Prewar China/Terry Sicular/228*

Reitsma, H. A., and J. M. G. Kleinpenning. *The Third World in Perspective/
Peter Dorner/229*

Sicular, Terry, ed. *Food Price Policy in Asia: A Comparative Study/
Howarth E. Bouis/230*

338.105
Am 35

Industry Members

and Representatives—1990/232

P 7688

American Journal of Agricultural Economics

Volume 73

• Number 2 •

May 1991

Articles

Area-Yield Crop Insurance Reconsidered/*Mario J. Miranda*/233

Scientific Principle and Practice in Agricultural Economics: A Historical Review/*Harold F. Breimyer*/243

Rent Seeking: The Potash Dispute between Canada and the United States/*Valerie J. Picketts, Andrew Schmitz, and Troy G. Schmitz*/255

Effects of Technological Change and Institutional Reform on Production Growth in Chinese Agriculture/*Shenggen Fan*/266

Farmer Behavior Under Risk of Failure/*William E. Foster and Gordon Rausser*/276

Valuing the Multidimensional Impacts of Environmental Policy: Theory and Methods/*John P. Hoehn*/289

Tax versus Quota Regulation: A Stochastic Model of the Fishery/*Robert A. Androkovich and Kenneth R. Stollery*/300

Endangered Species and the Safe Minimum Standard/*Richard C. Ready and Richard C. Bishop*/309

Demands for Local Public Sector Outputs in Rural and Urban Municipalities/*Melville L. McMilland and Joe Amoako-Tuffour*/313

Distributional Effects of Household Linkages/*G. Andrew Bernat, Jr., and Thomas G. Johnson*/326

Economic Impacts, Value Added, and Benefits in Regional Project Analysis/*Joel R. Hamilton, Norman K. Whittlesey, M. Henry Robison, and John Ellis*/334

Technical Change, Land Quality, and Income Distribution: A General Equilibrium Analysis/*Ian A. Coxhead and Peter G. Warr*/345

Modeling Agricultural Growth Multipliers/*Steven Haggblade, Jeffrey Hammer, and Peter Hazell*/361

Private Property Rights and Forest Preservation in Karnataka Western Ghats, India/*M. G. Bhat and R. G. Huffaker*/375

Impact Targets versus Discharge Standards in Agricultural Pollution Management/*John B. Braden, Robert S. Larson, and Edwin E. Herricks*/388

Tax Reform and Land-Using Sectors in the U.S. Economy: A General Equilibrium Analysis/*Roy Boyd and David H. Newman*/398

Consumer's Surplus Revisited/*J. S. Shonkwiler*/410

The Calculation of Research Benefits with Linear and Nonlinear Specifications of Demand and Supply Functions/*Jan P. Voon and Geoff W. Edwards*/415

Dairy Farm Efficiency Measurement Using Stochastic Frontiers and Neoclassical Duality/*Boris E. Bravo-Ureta and Laszlo Rieger*/421

Prospect Theory and Risk Preferences of Oregon Seed Producers/*Alan Collins, Wesley N. Musser, and Robert Mason*/429

Robustness of the Mean-Variance Model with Truncated Probability Distributions/*Steven D. Hanson and George W. Ladd*/436

Evaluating Robust Regression Techniques for Detrending Crop Yield Data with Nonnormal Errors/*Scott M. Swinton and Robert P. King*/446

Cointegration Tests and Spatial Price Linkages in Regional Cattle Markets/*Barry K. Goodwin and Ted C. Schroeder*/452

A Comparison of Video Cattle Auction and Regional Market Prices/*DeeVon Bailey, Monte C. Peterson, and B. Wade Brorsen*/465

Evaluating Prior Beliefs in a Demand System: The Case of Meats Demand in Canada/*James A. Chalfant, Richard S. Gray, and Kenneth J. White*/476

Ex Post Flexibility and Choice of Capacity for Loss Reduction/*L. Dean Hiebert*/491

The Application and Economic Interpretation of Selectivity Models/*Chung L. Huang, Robert Raunikaar, and Sukant Misra*/496

Erratum/502

Comments and Replies

The Impact of Increased LDC Food Production on LDC Food Imports: Comment/*Patrick J. Gormely*/503

The Impact of Increased Food Production on Less Developed Country Food Imports: Reply/*Alain de Janvry and Elisabeth Sadoulet*/506

Marketing Order Impacts on Farm-Retail Price Spreads: Comment/*Nicholas J. Powers*/507

Marketing Order Impacts on Farm-Retail Price Spreads: Reply/*Gary D. Thompson and Charles C. Lyon*/511

Rice in Asia: Is It Becoming an Inferior Good?: Comment/*Jikun Huang, Cristina C. David, and Bart Duff*/515

Rice in Asia: Is It Becoming a Commercial Good?: Comment/*Howarth E. Bouis*/522

Rice in Asia: Is It Becoming an Inferior Good?: Reply/*Shoichi Ito, E. Wesley F. Peterson, and Warren R. Grant*/528

The Efficiency of Alternative Policies for the EC's Common Agricultural Policy: Comment/*Alison Burrell*/533

The Efficiency of Alternative Policies for the EC's Common Agricultural Policy:
Reply/Harry de Gorter and Karl D. Meilke/535

Books Reviewed

Atkin, Michael. *Agricultural Commodity Markets: A Guide to Futures Trading*/Michael W. Woolverton/538

Becker, Tilman. *Die Weizenexportpolitik der Europäischen Gemeinschaft*/Stephen W. Hiemstra/538

Coombs, H. C. *The Return of Scarcity: Strategies for an Economic Future*/D. Lynn Forster/539

Demissie, Ejigou. *Small-Scale Agriculture in America: Race, Economics, and the Future*/Donald R. McDowell/540

Ferguson, Roy C., II. *Managing for Profit in Commercial Agriculture*/William J. Brown/541

Hancock, Graham. *Lords of Poverty: The Power, Prestige, and Corruption of the International Aid Business*/Eugene Jones/542

Harl, Neil E. *The Farm Debt Crisis of the 1980s*/Jerome M. Stam and Robert N. Collender/544

Morse, George W., ed. *The Retention and Expansion of Existing Businesses: Theory and Practice in Business Visitation Programs*/Gary B. Hansen/545

Moyer, H. Wayne, and Timothy E. Josling. *Agricultural Policy Reform: Politics and Process in the EC and the USA*/Robert D. Reinsel/546

Pasour, E. C., Jr. *Agriculture and the State: Market Processes and Bureaucracy*/Sandra S. Batie/547

Ph.D. Recipients by Subject, 1990/549

Ph.D. Recipients by Institutions, 1990/555

American Journal of Agricultural Economics

Volume 73

• Number 3 •

August 1991

Articles

Marketed Surplus Under Risk: Do Peasants Agree with Sandmo?/

Israel Finkelshtain and James A. Chalfant/557

Market Integration, Efficiency of Arbitrage, and Imperfect Competition:

Methodology and Application to U.S. Celery/*Richard J. Sexton,*

Catherine L. Kling, and Hoy F. Carman/568

An Empirical Analysis of Economic Performance Under the Marketing Order for

Raisins/*Ben C. French and Carole Frank Nuckton/581*

Are Farrowing Intentions Rational Forecasts?/*David E. Runkle/594*

The Welfare Analytics of Transaction Costs, Externalities, and Institutional

Choice/*Ronald C. Griffin/601*

Factor Demands in the U.S. Food-Manufacturing Industry/*Kuo S. Huang/615*

Government Intervention in Imperfectly Competitive Agricultural Input Markets/

Steve McCorriston and Ian M. Sheldon/621

Supply Impact of the Milk Diversion and Dairy Termination Programs/

Bruce L. Dixon, Dwi Susanto, and Calvin R. Berry/633

Determinants of Agricultural Economics Faculty Retirement/*Josef M. Broder,*

Fred C. White, and Teresa D. Taylor/641

Local Economic Conditions and Wage Labor Decisions of Farm and Rural

Nonfarm Couples/*J. G. Tokle and Wallace E. Huffman/652*

The Impact of Wage Differentials on Choosing to Work in Agriculture/

Jeffrey M. Perloff/671

Labor Contracting and a Theory of Contract Choice in California Agriculture/

Ann Vandeman, Elisabeth Sadoulet, and Alain de Janvry/681

The Welfare Effects of Targeted Export Subsidies: A General Equilibrium

Approach/*Mary Bohman, Colin A. Carter, and Jeffrey H. Dorfman/693*

The Costs of Indonesian Sugar Policy: A Policy Analysis Matrix Approach/

Gerald C. Nelson and Martin Panggabean/703

Education and Innovation Adoption in Agriculture: Evidence from Hybrid Rice in

China/*Justin Yifu Lin/713*

Incorporating Technological Change in Diffusion Models/*Mary K. Knudson/724*

A Bayesian Approach to Explaining Sequential Adoption of Components of a

Technological Package/*Howard D. Leathers and Melinda Smale/734*

On Testing the Structure of Risk Preferences in Agricultural Supply Analysis/
Rulon D. Pope and Richard E. Just/743

Factor Price Stabilization and the Competitive Firm/*S. Devadoss and
E. Kwan Choi/749*

Interactions between Agricultural and Resource Policy: The Importance of
Attitudes toward Risk /*Howard D. Leathers and John C. Quiggin/757*

Joint Risk Preference-Technology Estimation with a Primal System/*H. Alan Love
and Steven T. Buccola/765*

Construction of True Cost of Food Indexes from Estimated Engel Curves/
William Noel Blisard and James R. Blaylock/775

Wife's Employment, Food Expenditures, and Apparent Nutrient Intake: Evidence
from Canada/*Susan Horton and Cathy Campbell/784*

Long-Run versus Short-Run Planning Horizons and the Rangeland Stocking Rate
Decision/*L. Allen Torell, Kenneth S. Lyon, and E. Bruce Godfrey/795*

Testing for Consistent Aggregation/*Robert G. Chambers and Rulon D. Pope/808*

On Nonlinear Dynamics: The Case of the Pork Cycle/*Jean-Paul Chavas and
Matthew T. Holt/819*

State-Space Modeling of Cyclical Supply, Seasonal Demand, and Agricultural
Inventories/*Jeffrey H. Dorfman and Arthur Haverner/829*

Perennial Crop Supply Response: A Kalman Filter Approach/*Keith C. Knapp and
Kazin Konyar/841*

Measuring the Potential Contribution of Plant Breeding to Crop Yields: Flue-
Cured Tobacco, 1954-87/*Bruce A. Babcock and William E. Foster/850*

Using Count Data Models in Travel Cost Analysis with Aggregate Data/
Daniel M. Hellerstein/860

Erratum/867

Comments and Replies

A Model of Production with Supply Management for the Canadian Agricultural
Sector: Comment/*Carl H. Nelson/868*

A Model of Production with Supply Management for the Canadian Agricultural
Sector: Reply/*Giancarlo Moschini/871*

Elasticities in AIDS Models: A Clarification and Extension/*Richard Green and
Julian M. Alston/874*

Proceedings

Technology Policy and Agriculture

Emerging Issues in the Allocation of Public Agricultural Research Funds/
George B. Frisvold/876

Plant Variety Protection, Private Funding, and Public Sector Research Priorities/
Mary K. Knudson and Carl E. Pray/882

Inventions Intended for Use in Agriculture and Related Industries: International
Comparisons/*Robert E. Evenson/887*

International Technology Transfer: Private Channels and Public Welfare/
Margot Anderson and Bruce A. Larson/892

Technology Policy and Agriculture: Discussion/*John Reilly and Roger Conway/*
898

Technology Policy and Agriculture: Discussion/*Marie E. Walsh/901*

Technology Policy and Agriculture: Discussion/*James F. Oehmke/903*

The 1990 Farm Bill and the GATT Outcome

The Effect of the 1990 Farm Bill on Agricultural Trade/*George E. Rossmiller*
and *Rachel A. Nugent/905*

The 1990 Farm Bill and the Uruguay Round/*C. Ford Runge/909*

Impacts of the 1990 Farm Bill on Consumers/*Carol S. Kramer/913*

Farm Policy Reform and the Environment/*John M. Antle/917*

Special Interests and the 1990 Farm Bill: Discussion/*B. Delworth Gardner/922*

The 1990 Farm Bill and the Uruguay Round: Discussion/*Daniel A. Sumner/924*

Data Needs for the Assessment of Environmental Quality and Food Safety

Data Needs to Assess Environmental Quality Issues Related to Agriculture and
Rural Areas/*Jerald J. Fletcher and Tim T. Phipps/926*

Data Needs to Address Economic Issues in Food Safety/*Tanya Roberts and*
David Smallwood/933

Data Needs to Assess Environmental Quality Issues Related to Agriculture and
Rural Areas: Discussion/*David E. Ervin/943*

Data Needs to Address Economic Issues in Food Safety: Discussion/
Julie A. Caswell/945

International Capital Markets and Development Funds for Agriculture

International Capital Markets, Development Performance, and Lessons for the
Future/*Uma Lele/947*

Investments to Transfer Poultry Production to Developing Countries/*Gary Vocke/*
951

International Capital Markets and Interest Groups/*Terry Roe/955*

International Capital Markets and Development Funds for Agriculture: Discussion/
Ralph W. Cummings/963

Books Reviewed

Anderson, Kym. *Changing Comparative Advantages in China: Effects on Food, Feed and Fibre Markets*/Colin A. Carter/1965

Bourman, F. J. A. *Small, Short and Unsecured: Informal Rural Finance in India*/Mario N. Lamberte/1966

Dreze, Jean, and Amartya Sen. *Hunger and Public Action*/Dana G. Dalrymple/1967

Eicher, Carl M., and John M. Staatz, eds. *Agricultural Development in the Third World*/Ronald L. Tinnermeier/1968

Fleischer, Beverly. *Agricultural Risk Management*; Smidts, A. *Decision Making Under Risk* /Walter L. Fishel/1969

Gray, Kenneth R., ed. *Soviet Agriculture: Comparative Perspectives*; Wädekin, Karl-Eugen, ed. *Communist Agriculture: Farming in the Soviet Union and Eastern Europe*/Joseph Havlicek, Jr./1970

Jackson, John H. *Restructuring the GATT System* /Bradley J. McDonald/1971

Lobao, Linda M. *Locality and Inequality: Farm and Industry Structure and Socioeconomic Conditions*/Jerry R. Skees/1972

Longworth, John W., ed. *China's Rural Development Miracle, with International Comparisons* /Richard L. Meyer/1973

Pearce, David. *Sustainable Development: Economics and Environment in the Third World* /Daniel W. Bromley/1975

van der Meer, Cornelis L. J., and Saburo Yamada. *Japanese Agriculture, A Comparative Economic Analysis* /Frank Hsiao/1976

Formally Testing the Present Value Model of Farmland Prices

Barry Falk

This paper applies recent statistical developments in the study of stock market price movements to study the plausibility of the constant expected returns version of the present value model as an explanation of farmland prices. Using Iowa farmland price and rent data over the 1921–86 sample period, formal test results indicate that, although farmland price and rent movements are highly correlated, price movements are not consistent with the implications of this model. There appear to be persistent predictable excess positive and/or negative returns in the Iowa farmland market.

Key words: efficient market hypothesis, farmland prices, present value model.

Farmland prices in the United States have undergone occasional periods of sustained booms followed by sustained busts. The most recent example was the boom of the 1970s and the bust of the 1980s. Melichar (1984) has identified four other price cycles of this sort over the last two hundred years. The recent major cycle in farmland prices has generated considerable interest in the determination of farmland prices. Papers by Alston, Melichar (1979), Burt, and Featherstone and Baker are especially relevant to this study.

The financial economics literature also has displayed a surge of interest in evaluating the extent to which financial asset prices can be explained in terms of rational evaluations of current and expected asset income. This literature began with studies by Shiller and by Leroy and Porter. The current state of this literature is summarized by DeBondt and Thaler and by West (1988a).

The fundamental economic questions addressed in these two lines of literature are essentially the same. First, how much of an asset's price movement is caused by factors which determine the fundamental value of the asset,

i.e., current and expected future returns and discount rates? Second, do these factors account for asset price movements in a rational manner? In contrast to the existing studies of farmland prices, the financial economics literature has developed new statistical tests (such as Shiller's variance bounds tests) designed especially to help answer these kinds of questions; and it has utilized recent developments in the study of non-stationary time series, such as the application of unit root and cointegration theory.

The purpose of this paper is to apply statistical developments in the study of stock market price movements to study farmland prices and formally test the validity of the constant discount rate version of the present value model of farmland prices. A strategy developed by Campbell and Shiller (1987) will be applied to study the behavior of annual Iowa farmland prices over the 1921–86 sample period. The Iowa farmland market is used for several reasons. First, relatively long time series on Iowa farmland prices and rents are available. Second, Iowa farmland is relatively homogenous, primarily suited for growing corn and soybeans, and its value has not been strongly influenced by non-agricultural uses. Finally, an active cash rental market provides readily available cash rent data for Iowa farmland.

The remainder of the paper provides a review of the literature on farmland prices, a description of Campbell and Shiller's procedures, a description of the data, the empirical results and a discussion of their implications.

Barry Falk is an associate professor of economics, Iowa State University.

Journal Paper No. J-14034 of the Iowa Agriculture and Home Economics Experiment Station Project No. 2911.

The author gratefully acknowledges Michael Duffy's help in obtaining the data, the research assistance of Cheol-Soo Park, and helpful comments from anonymous referees and Robert Shiller.

Review of Related Literature

The modern study of farmland price determination goes back at least to the 1960s with the efforts of Herdt and Cochrane, Reynolds and Timmons, and Tweeten and Martin to model land prices in a simultaneous equations framework; but the inability of these models to explain subsequent land prices (as shown by Pope et al.) discouraged further efforts along these lines. Much of the problem in this kind of approach involves the inelasticity of farmland quantity, so that attempts to identify a classic supply equation in this market may not be appropriate. Since the 1970s, attention turned to the role of demand forces in explaining land price movements.

Klinefelter, and Castle and Hoch attempted to explain land prices in terms of expected net rents and expected capital gains as though these two components of land price are distinct from one another. Although it is possible that capital gains are driven by speculative forces or by forces unrelated to rents (i.e., by "sunspots"), capital gains can also result from a growing stream of net rents (Melichar 1979). More recent literature has focused on the behavior of land prices in a manner consistent with the conventional present value formulation.

Consider the conventional present value formulation of farmland price determination:

$$(1) \quad Y_t = \beta \sum_{s=0}^{\infty} \beta^s E[y_{t+s}|I_t],$$

where Y_t is the real value per acre of homogeneous farmland at the start of period t , y_t is the real net rent per acre of farmland in period t (paid at the end of period t), and β is a discount factor which is equal to $1/(1+r)$ where r is the real interest rate, so that $0 < \beta < 1$. Expressions of the form $E[X_{t+s}|I_t]$ denote the market's expectation of X_{t+s} formed at the start of time t on the basis of its information set at that time. This model assumes a constant real interest rate and ignores the differential tax treatment of capital gains and rental income. It can be derived as the stable forward solution to the following difference equation which is a form of the efficient markets hypothesis:

$$(2) \quad Y_t = E[Y_{t+1} + y_t|I_t]/(1+r).$$

The fact that (1) can be derived as the solution to (2), solving (2) forward recursively, makes clear the point that capital gains can be ex-

plained in theory as the capitalization of expected future rents.

Melichar (1979) and Alston assume that net rents are expected to grow at the constant rate g so that $E[y_{t+s}|I_t] = (1+g)^s y_t$ for all t . In this case, the value of Y_t implied by (1) simplifies to $Y_t = (1/(r-g))y_t$, provided that r is larger than g . According to this model, land prices should grow at the same rate as rents grow. Using U.S. aggregate data since 1950 (Melichar) and data from eight midwestern states since 1960 (Alston), these authors compare average annual growth rates of real land prices and real returns to land ownership over various sample periods and find that these averages are close to one another. They conclude that most of the variation in farmland prices over time can be attributed to changes in net rental income.

Burt assumes that, although (1) may provide a reasonable description of the long-run or steady-state equilibrium behavior of farmland price determination, it is less likely to be an appropriate representation of the short-run behavior of farmland prices. However, net rents may still be the primary factors determining land prices. To quantify this view Burt assumes that Y_t is determined according to the distributed lag model,

$$(3) \quad Y_t = (\alpha y_t^{b_0} y_{t-1}^{b_1} \dots) u_t,$$

where $\log u_t$ is a zero-mean stationary process, α is the inverse of the real interest rate, and b_0, b_1, \dots are fixed parameters which sum to unity.¹ He estimates a restricted version of (3) and then shows according to a variety of descriptive criteria that it can provide a good in-sample fit to Illinois annual farmland price data over the period 1960-84 and that it has good out-of-sample predictive power (treating out-of-sample rents as known *ex ante*). He concludes (p. 25) that his results "strongly support the view that rents are the underlying source of value, and there is little evidence that farmland prices are driven by the same speculative forces as those for nonincome-earning assets such as precious metals and stones." He also concludes, however, that there are aspects of the relationship between land prices and rents which are not consistent with the present value mode.

In contrast to Alston, Burt, and Melichar (1979), Featherstone and Baker conclude that

¹ Burt assumed that in the long run rents converge to the constant y^* . Because the unconditional expectation of u_t is equal to one and the b 's sum to one, it follows from (3) that Y_t will converge to αy^* , which is consistent with (1) under the assumption that y_t is equal to y^* for all t .

net rents cannot explain a substantial portion of farmland price movements, suggesting that there may be purely speculative forces in farmland price determination. Furthermore, to the extent that farmland prices do respond to unanticipated changes in current and expected future rents, the responses are too large and protracted to be consistent with a present value formulation. These conclusions are based on an application of Sims's innovation accounting method to an estimated vector autoregression of U.S. annual real farm asset values, real net returns, and real interest rates.

In summary, a review of the recent literature indicates that a variety of strategies have been used to evaluate the plausibility of the present value model of farmland prices with mixed conclusions. However, none of the strategies incorporate a formal test of the restrictions that the model imposes on the data. The main purpose of this paper is to fill that gap.

A Strategy for Testing the Present Value Model of Farmland Prices

The purpose of this section is to explain the statistical procedures that will be used to evaluate whether Iowa farmland prices evolve according to the present value relationship. In contrast to the literature reviewed above, these procedures will lead to formal tests of the model. The strategy is fully described by Campbell and Shiller (1987) and thus will be only briefly reviewed here.

We begin with the assumption that the present value model (1) is a valid representation of the relationship between real farmland prices at the start of year t and current and expected future real net rents, where current rent accrues to the landowner at the end of period t . The model assumes that the discount rate β and, hence, the real interest rate r are constants.² The market's expectations are assumed rational so that $E[y_{t+s}|I_t]$

is interpreted as the mathematical conditional expectation of y_{t+s} given the information set I_t . The only restriction imposed on this information set is that it contains at least Y_t, Y_{t-1}, \dots and y_t, y_{t-1}, \dots .³ Let H_t denote this subset of I_t .

Real net rent time series show a tendency to grow over time; therefore, real net rents are nonstationary in their mean. One common modeling approach is to assume that the process is stationary around a nonstochastic linear trend. An alternative assumption is that the process is difference stationary, i.e., the first differences of the process form a stationary process. The choice of assumption has important statistical and economic implications.⁴ For example, an unanticipated change in a trend stationary process has only a temporary effect on the process, whereas an unanticipated change in a difference stationary process has a permanent effect. Assume for now that net rents evolve as a difference stationary process.

If the present value model is correct and if net rents evolve as a difference stationary process, then it can be shown (Campbell and Shiller 1987) that Y_t is also a difference stationary process. One battery of tests of the present value model will consist of Dickey-Fuller tests to determine whether the pattern of stationarity between land prices and rents is consistent with this implication of the present value model.

Next, define a new variable S_t called the spread which is the linear combination of Y_t and y_t given by

$$(4) \quad S_t = Y_t - \theta y_t,$$

where θ is determined by the discount factor according to $\theta = \beta/(1 - \beta)$; thus, it equals the reciprocal of the real interest rate. Campbell and Shiller show that if the present value model is correct, then the spread can be interpreted as the rational forecast of the present value of all future changes in net rents.

³ This restriction requires an assumption that cash rent payments in year t are known at the start of the year which seems reasonable if these payments are based upon contracts written at the start of that year. Campbell and Shiller (1987) show how to modify their procedures if these payments are not determined until later in the year. All of the empirical analysis in this paper was also executed in terms of that modified procedure. The results were not sensitive to these modifications.

⁴ Whether economic time series are better characterized as difference stationary processes or as trend stationary processes has been an important issue in studying the relationship between stock market prices and dividends (West 1988a). It is also at the core of the current debate in macroeconomics over whether real business cycle models or monetary business cycle models offer a more plausible explanation of the business cycle. At this point the prevailing view appears to favor the assumption of difference stationarity.

² Burt also assumed that the discount rate is constant over time. Although Featherstone and Baker incorporated observed real interest rates in their VAR analysis to account for a time-varying discount rate, changes in the real interest rate had only a marginal effect on the time path of land prices. While it is possible to allow for a time-varying discount rate (e.g., Campbell and Shiller 1988), this requires additional less appealing restrictions on the model's structure. Furthermore, the constant discount rate version of the model is more convenient to work with analytically and empirically. It should not be abandoned without compelling empirical arguments. The absence of a formal test of the constant discount rate version of the present value model of farmland prices, prior to this paper, indicates that such arguments do not yet exist.

The present value model imposes testable cross-equation restrictions on the vector autoregressive representation of the change in $y_t(\Delta y_t)$ and the spread S_t . Specifically, if Δy_t is assumed stationary then the present value model implies that S_t will be stationary too. Then, under general side conditions, Δy_t and S_t will have a vector autoregressive representation of the form

$$(5a) \quad \Delta y_t = a(L)\Delta y_{t-1} + b(L)S_{t-1} + u_{1t}$$

$$(5b) \quad S_t = c(L)\Delta y_{t-1} + d(L)S_{t-1} + u_{2t},$$

where $a(L)$, $b(L)$, $c(L)$, and $d(L)$ are p th-order polynomials in the lag operator L , i.e., $a(L) = a_0 + a_1L + a_2L^2 + \dots + a_pL^p$, where $L^p x_t \equiv x_{t-p}$, and u_{1t} and u_{2t} are zero-mean, constant variance, and serially uncorrelated processes which are mutually uncorrelated at all leads and lags, except that $E[u_{1t}u_{2t}]$ may be nonzero. The present value model and the stationarity of Δy_t imply that this VAR will be characterized by S_t Granger-causing the Δy_t process, so that some of the elements of $b(L)$ will be nonzero, and by the following cross-equation restrictions: $c_i = -\theta a_i$, $i = 1, 2, \dots, p$; $d_1 = (1/\beta) - \theta b_1$; and $d_i = -\theta b_i$, $i = 2, \dots, p$.

The restriction that S_t Granger-causes the Δy_t process means that, given past changes in net rents, past values of the spread add useful information in forecasting future changes in rents. Recall that the spread can be interpreted as an optimal forecast of a weighted average of future changes in net rents based upon the market's complete information set. Any information in this set beyond lagged values of y_t itself that is useful in forecasting future changes in y_t will be reflected in the VAR by a nonzero $b(L)$. This situation is formally proven by Campbell and Shiller (1987, footnote 7).

To interpret the cross-equation restrictions define a new variable, x_t , according to

$$(6) \quad x_t \equiv Y_t + y_{t-1} - (1+r)Y_{t-1}.$$

Thus, according to (2), x_t measures the *ex post* excess return to holding land in period $t - 1$. According to the present value model (1), it can be shown that another representation of x_t is

$$(7) \quad x_t = Y_t - E[Y_t|I_{t-1}].$$

This follows because the present value model is constructed to rule out *ex ante* excess returns and the only component of x_t which is assumed to be unknown at the start of period $t - 1$ is the end-of-the-period price of land, Y_t . Thus, x_t can also be interpreted as the unpredictable com-

ponent in the price of land in period t based upon the market's complete information set in period $t - 1$, I_{t-1} . Because this information set contains, by assumption, at least the set of past prices and returns, H_{t-1} , it follows by the law of iterated expectations that $E[x_t|H_{t-1}] = 0$ for all t . In other words, according to the present value model, current excess returns ought to be unpredictable based upon past prices and rents. The cross-equation restrictions on the VAR coefficients account for this implication of the model.

In summary, if land prices evolve according to the present value model (1) where market expectations are formed rationally on the basis of an information set which contains at least past prices and rents and if rents evolve as a difference stationary process, then the model has testable implications of two types. First, it implies that land prices also evolve as a difference stationary process. Second, the theory imposes Granger-causality and cross-equation restrictions on the vector autoregressive representation of the spread and the changes in rents.

The Data

In this paper Y_t is measured as the estimated average value of an acre of Iowa farmland in year t and y_t is measured as the estimated average annual cash rent per acre of Iowa farmland in year t . Each series is divided by the corresponding year's consumer price index (1967 = 100). The sample period is 1921–86.

The time series of average annual prices per acre of Iowa farmland was constructed by splicing the U.S. Department of Agriculture's (USDA) farm real estate value, by state, series (1921–49) with Iowa State University Extension Service's Iowa land value series (1950–86). The latter series is considered to be a more reliable measure of average Iowa land prices since 1950. The differences between the two series since 1950 are not large enough to substantially affect the results of this study. The USDA's series is described more fully in Barnard and Hexum. The Iowa Land Value Survey data are described more fully in the Iowa State University Extension Service's FM-1825 publications.

The theoretical model assumes that Y_t is the price at the beginning of the period. It is not possible to associate a particular part of the year to which the USDA price measure most closely corresponds. The Iowa Land Value Survey reports end-of-period prices (i.e., November of year t). Although the results reported in this paper

treat all of these prices as beginning-of-the-year prices, the empirical procedures were also done using the price reported in year $t - 1$ (divided by the price index for year t) to form the measure of Y_t . The latter results (available from the author) are less consistent with the present value model.

The cash rent time series used to form y_t are average dollar rents paid per acre for the rental of whole Iowa farms (i.e., land and buildings). These data are estimates produced by the Iowa Agricultural Statistics Service (formerly known as the Iowa Crop and Livestock Reporting Service). They are published by the USDA's Economic Research Service. Using cash rents may have several shortcomings. Several types of farmland leasing contracts are common in Iowa: crop-share leases, cash leases, and combinations of the two. Perhaps a systematic relationship between the quality of land and the nature of the leasing arrangement biases the cash rent data as a measure of return per acre. Second, factors other than the cash rent influence the landowner's net return, such as owner-borne maintenance costs, property taxes, and insurance premiums. Nevertheless, the active cash rent market in Iowa suggests that the use of cash rents to index returns may be less of a problem than it would in other parts of the country.

The cash rent and cash price series were converted to real rent and price series, respectively, by dividing each series by the CPI (all items, 1967 = 100) for that year. The CPI data were collected from the U.S. Commerce Department's publication *Historical Statistics of the United States: Colonial Times to 1970* (1921–70) and from various issues of the Commerce Department's *Business Statistics*. The choice of an appropriate deflator is problematic. The CPI was selected because it is the purchasing power of consumption goods that motivates portfolio allocation choices. It is unlikely that the results will be sensitive to the use of other deflators.

Figure 1 is a time series plot of Y_t and y_t , where the latter is weighted by a constant to provide a common scale. Both price and rent are characterized by positive long-run trends which were interrupted by the major declines during the 1920s and early 1930s and again in the late 1970s and early 1980s. The similarity in the general shapes of the two series is confirmed by a sample cross-correlation statistic of 0.90. However, casual observation suggests that prices fall faster than rents when both are falling and prices rise faster than rents when both are rising. In other words, price movements appear to be more volatile than

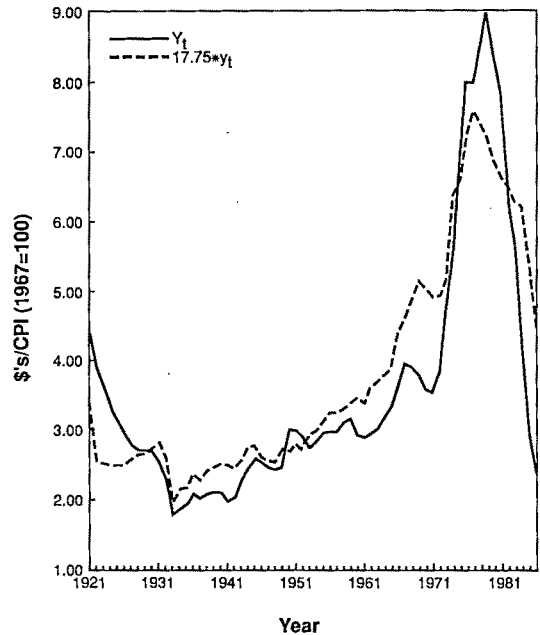


Figure 1. Price (Y_t) and weighted rent ($17.75*y_t$)

rent movements. This paper will formally test whether prices are excessively volatile relative to the predictions of the present value model.

Test Results

Unit Root Tests

The test strategy discussion indicated that the tests are not valid unless real rents evolve as a difference stationary process. In this case, the theory implies that land prices will also evolve as a difference stationary process. Thus, the first test is whether land prices and rents evolve as difference stationary processes rather than trend stationary processes.

Consider the following representations of rents and prices, respectively,

$$(7) \quad \Delta y_t = \alpha + \delta t + \phi y_{t-1} + u_t, \text{ and}$$

$$(8) \quad \Delta Y_t = \alpha' + \delta' t + \phi' Y_{t-1} + v_t,$$

where u_t and v_t are zero-mean, constant variance, and serially uncorrelated disturbance processes. Under the null hypothesis that $y_t(Y_t)$ is difference stationary the parameters δ and ϕ (δ' and ϕ') are jointly equal to zero. Under trend stationarity, δ (δ') is unrestricted and ϕ (ϕ') is between zero and negative one. The Dickey-Fuller test of the null hypothesis is based upon

a comparison of the standard t -statistic obtained from a regression of Δy_t on a constant, t , and y_{t-1} (and of ΔY_t on a constant, t , and Y_{t-1}) and the critical values of this statistic under the null hypothesis. The distribution of the t -statistic under the null hypothesis is nonstandard, and so the critical values cannot be extracted from the usual tables. The critical values for the non-standard distribution are available in Fuller (p. 373).

Table 1 shows the Dickey-Fuller test results. The null hypothesis that land prices and rents are stationary in their respective first differences cannot be rejected.⁵ Thus, the crucial precondition of the testing strategy and one of the subsequent implications of the present value model are assumed to be satisfied.

Testing the VAR Restrictions

The next step is to test the restrictions that the present value model imposes on the VAR representation of Δy_t and S_t . The process S_t is not directly observable since according to (4) it depends upon θ , the reciprocal of the unobservable constant real interest rate appropriate for discounting future income. However, a consistent estimator of θ is the regression coefficient on y_t in a regression of Y_t on y_t and a constant.⁶ This

estimate of θ is 17.75, which implies an annual real interest rate of 5.6%.⁷ Based upon this estimate of θ we can construct a consistent estimator of S_t according to (4).

The unrestricted VAR representation of Δy_t and S_t was assumed to have a third-order lag structure based upon the results of Sims' (1980) modified likelihood ratio test. Table 2 contains a summary of the VAR, which was estimated by OLS, including the results of tests of the restrictions implied by the present value model. The R^2 of the Δy equation implies that over 50% of the change in real cash rents is predictable on the basis of its own past changes and the past history of the spread. The Granger-causality test cannot reject the hypothesis that the spread Granger-causes changes in rents, which is an implication of the present value model provided that the market uses information other than past values of y_t to forecast current and future changes in y_t .

The cross-equation linear restrictions implied by the present value model also are tested. These restrictions correspond to the restriction that excess returns in the Iowa farmland market are not predictable and, hence, that systematic profit opportunities do not typically exist. Wald tests were executed in RATS based upon White's heteroskedasticity-consistent covariance matrix estimator. This estimator accounts for the difference between the variances of the two disturbance processes in the VAR as well as the contemporaneous correlation between those two processes. The model was estimated without restricting the mean of either Δy_t or S_t . However, the present value model implies that the unconditional mean of the spread is zero. Therefore the cross-equation restrictions were tested without restricting the mean of S_t , and then these restrictions were tested along with this zero restriction. Both versions of the test reject the cross-equation restrictions implied by the present value model at the 1% level of significance.⁸

⁵ A modified version of the Dickey-Fuller test proposed by Said and Dickey was applied to account for possible serial correlation in the disturbances in (7) and (8). A version of the test developed by Phillips and Perron was applied to account for possible serial correlation or heteroskedasticity in these disturbances. The results consistently support the null hypothesis of difference stationarity. The possibility of a second unit root in each series was also considered and rejected. This is important because Diba and Grossman have shown that if y_t has exactly one unit root then the presence of at least two unit roots in Y_t is a necessary condition for deviations from the present value model because of a rational bubble. Further discussion and results pertaining to the unit root issue are available in Falk.

⁶ This follows because one way to interpret the implication of the theory is that Y_t and y_t are cointegrated processes of order (1, 1). See Campbell and Shiller (1987) or, for a more complete discussion of cointegrated processes, see Engle and Granger.

Table 1. Dickey-Fuller Unit Root Tests

Unrestricted model: $\Delta z_t = \alpha + \delta t + \gamma z_{t-1} + u_t$

$$H_0: \delta = \gamma = 0$$

(z_t is difference stationary)

z_t	τ
y_t	-1.25
Y_t	-2.50

Note: τ is the " t -statistic" on the regression coefficient on z_{t-1} in the unrestricted model. H_0 is rejected at the 10% level of significance if $\tau < -3.18$. See the third section of Fuller's table 8.5.2.

Further Analysis

It is possible that the statistical rejection of the implications of the present value model is not of economic importance. That is, the statistical

⁷ The mean annual rate of return in this market, measured as the sample mean of $(Y_{t+1} - Y_t + y_t)/Y_t$, was 5.7%.

⁸ The distribution of the Wald statistic under the null hypothesis is based upon asymptotic distribution theory. Campbell and Shiller's (1989) Monte Carlo experiments on the small-sample properties of a similar test statistic indicate a slight tendency to reject the null too often for a given test size.

Table 2. Testing the Restrictions of the VAR Representation of Δy_t and S_t

Sample Period: 1921–1986, annual observations

Lag Length = 3

Granger-Causality Tests—

 Δy equation $R^2 = .522$; S Granger-causes Δy at the .00001 level. S equation $R^2 = .825$; Δy Granger-causes S at the .018 level.

Wald Tests of the Cross-Equation Restrictions—

without the mean restriction: $\chi^2(6) = 38.27$, p -value = .99E-06with the mean restriction: $\chi^2(7) = 39.38$, p -value = .17E-05.

failure of the present value model implies that the absence of systematic excess profit opportunities in the Iowa land market is quite unlikely. However, the magnitude of the profit opportunities may be small enough to be disregarded. In this section an argument is made that the statistical results reported above are the result of economically important deviations from the present value model.

Let S'_t denote the expected present value of all future changes in rent conditional upon current and past prices and rents, multiplied by the constant θ :

$$(9) \quad S'_t = \theta \sum_{i=1}^{\infty} \beta^i E[\Delta y_{t+i} | H_t],$$

where the information set H_t contains current and past values of land prices and rents. Based upon the unrestricted VAR representation of Δy_t and S_t given by (5) it can be shown (Campbell and Shiller, 1987, p. 1068) that

$$(10) \quad S'_t = \theta h' \beta A (I - \beta A)^{-1} z_t,$$

where h' is the $1 \times 2p$ row vector $[1 \ 0 \ \dots \ 0]$; I is the $2p \times 2p$ identity matrix; $z_t = [\Delta y_t, \dots, \Delta y_{t-p+1}, S_t, \dots, S_{t-p+1}]'$; and A is the $2p \times 2p$ companion matrix of the VAR, i.e.,

$$A = \begin{bmatrix} a_1 & a_2 & a_3 & \dots & a_{p-1} & a_p & \dots & b_1 & b_2 & b_3 & \dots & b_{p-1} & b_p \\ 1 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ c_1 & c_2 & c_3 & \dots & c_{p-1} & c_p & \dots & d_1 & d_2 & d_3 & \dots & d_{p-1} & d_p \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 1 & 0 \end{bmatrix}$$

Thus, S'_t can be estimated based upon the previous estimates of θ (and, hence, β) and the VAR coefficients.

The observed spread, S_t , and the theoretical spread, S'_t , are related according to

$$(11) \quad S'_t - S_t = \sum_{i=1}^{\infty} \beta^i E[x_{t+i} | H_t],$$

where x_t was defined previously as the excess return in period t to land ownership.⁹ Thus, the difference between the theoretical spread and the actual spread measures the expected present value of predictable excess future returns. According to the present value model, predictable excess returns should be zero, i.e., the actual and theoretical spreads should be equivalent. Therefore, sampling error aside, observed deviations of the actual spread from the theoretical spread indicate persistent predictable excess positive and/or negative returns in the Iowa farmland market.

Figure 2 indicates a strong degree of negative correlation between the time-series plots of S_t and S'_t . The sample correlation between these two time series is $-.86$ with a numerically estimated standard error of $.07$. The extreme negative correlation is too large to be attributable to sampling error or to measurement errors associated with the use of cash rents to measure net returns to land ownership.

It is interesting to compare these results with those of Campbell and Shiller in their study of stock market prices. They also rejected the cross-equation restrictions on the VAR. When they used the sample mean annual rate of return on stock ownership (8.2%), they computed a correlation between the theoretical and actual spreads of $-.46$. Thus, the results for the U.S. stock market and the Iowa farmland market are consistent with one another and suggest the presence of some common underlying tendencies. This consistency also helps to confirm the value

⁹ Equation (11) can be derived from equations (4), (6), and (10).

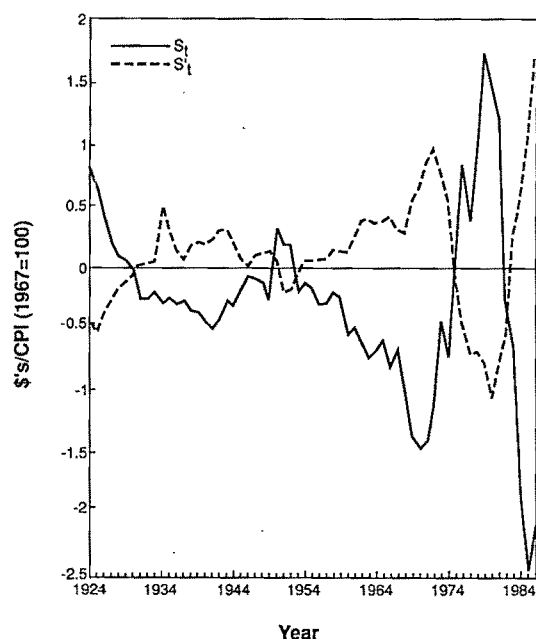


Figure 2. Theoretical spread (S'_t) and actual spread (S_t)

of applying this analytical approach to the study of farmland prices.

On the basis of figure 2 it appears that the statistical rejection of the present value model is attributable to sustained periods of positive (or negative) expected excess returns to holding Iowa farmland. Furthermore, the sign and magnitude of these expected excess returns vary directly with the sign and magnitude of the theoretical spread (and inversely with the sign and magnitude of the actual spread). Loosely speaking, Iowa farmland prices tend to be unusually high (low) when the present values of expected future changes in rents are unusually low (high).

One possible explanation of these features of the data is that traders in the Iowa farmland market act in a myopic manner. Suppose that the changes in rents are positively autocorrelated and a sequence of decreases in real rents (relative to the mean change) occurs. The stationarity of these first differences means that the sequences of decreases will subsequently be offset by a sequence of increases before stabilizing at their mean value. If, however, traders ignore the tendency of runs of decreases to be offset by runs of increases and instead assume that recent changes are permanent or will be exacerbated over time, then the price of land will be driven downward. Although the theoretical spread will

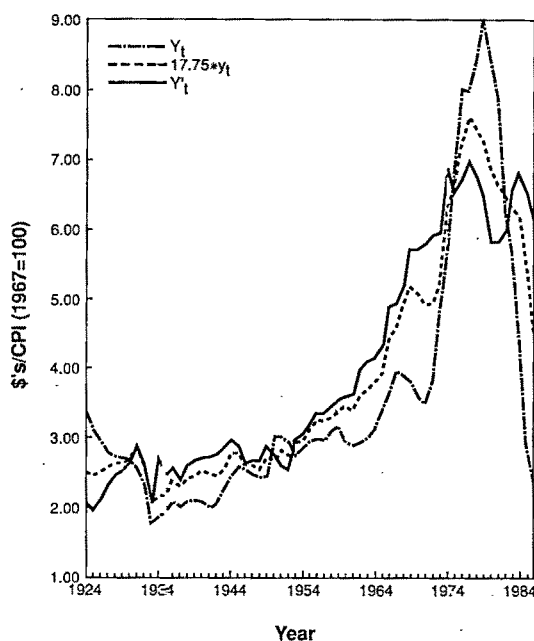


Figure 3. Price (Y_t), weighted rent ($17.75*y_t$), and rational price (Y'_t)

be positive due to the transitory fall in rents, the actual spread will be negative because of the erroneous perception that the fall is permanent.

Another view is offered by comparing the behavior of the actual price, the *ex ante* rational price, and actual rents.¹⁰ Figure 3 contains a plot of the actual real price (Y_t), the *ex ante* rational price (Y'_t) derived according to $Y'_t = S'_t + \theta y_t$, and real rents weighted by the constant θ .¹¹ Notice that Y'_t tends to move less than proportionally with respect to changes in y_t , while Y_t tends to move more than proportionally with respect to such changes. In other words, Iowa farmland prices show a tendency to overreact to movements in rents. These tendencies are especially apparent during the most volatile periods of the sample period. They suggest that θy_t will tend to lie between Y_t and Y'_t , and, therefore, S_t will be negatively correlated with S'_t .

Concluding Comments

Recent studies of farmland prices have suggested that net rents are the primary determinant

¹⁰ This explanation and the use of figure 3 to support it were suggested to me by Robert Shiller.

¹¹ The rational price would prevail according to (4) if the theoretical spread (S'_t) and the actual spread (S_t) are equal to one another, i.e., in the absence of exploitable profit opportunities.

of farmland prices (Alston, Burt, and Melichar) but that these prices tend to overreact to recent changes in rents (Featherstone and Baker). This paper has utilized statistical techniques developed in the financial economics literature to study the relationship between financial asset prices and returns in order to assess these conclusions. Using Iowa farmland price and rent data over the 1921–86 sample period, the results indicate that although farmland price and rent movements are highly correlated, price movements are much more volatile than rent movements. A formal test of the cross-equation restrictions imposed by the present value model of farmland price determination failed to support these restrictions and, therefore, the model itself. Subsequent analysis suggested that the model's statistical failure is attributable to economic factors rather than to measurement and sampling error.

One possible explanation of the model's failure is that the Iowa farmland market is characterized by rational bubbles. A rational bubble reflects a tendency for price to deviate from its fundamental value in a nonstationary manner as a result of self-fulfilling beliefs that the price depends on a variable (or a group of variables) that may be intrinsically irrelevant with respect to the asset's fundamental value (e.g., Diba and Grossman). The presence of such bubbles does not appear to characterize the U.S. stock market, and Falk has argued that they seem unlikely in Iowa farmland prices as well.

Another possible explanation that is consistent with rational investor behavior is that the discount rate in the present value model is time varying. Shiller, West (1988b), and Campbell and Shiller (1988) concluded that this explanation cannot reconcile the behavior of stock market prices and dividends with the present value model. Nonetheless, further research could explore this issue for farmland valuation.

Perhaps the most promising direction for future research into the determination of farmland prices is the study of nontraditional, or fad, models. In these models (e.g., Summers), prices tend to overreact to movements in current and recent returns. Although price has a tendency to return ultimately to its fundamental value (in contrast to a rational bubble), it may be years before this return occurs. In the meantime, systematic profit opportunities exist in the market although they may be difficult to detect and exploit.

[Received October 1988; final revision received March 1990.]

References

- Alston, J. "An Analysis of Growth of U.S. Farmland Prices: 1963–1982." *Amer. J. Agr. Econ.* 68(1986):1–9.
- Barnard, C. H., and R. Hexum. *Major Statistical Series of the U.S. Department of Agriculture: Land Values and Land Use*. Washington DC: U.S. Department of Agriculture, Econ. Res. Serv. Agr. Handbook No. 671, vol. 6, 1988.
- Burt, O. R. "Econometric Modeling of the Capitalization Formula for Farmland Prices." *Amer. J. Agr. Econ.* 68(1986):10–26.
- Campbell, J., and R. Shiller. "Cointegration and Tests of Present Value Models." *J. Polit. Econ.* 95(1987):1062–88.
- . "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors." *Rev. Finan. Stud.* 1(1988):195–228.
- . "The Dividend Ratio Model and Small Sample Bias: A Monte Carlo Study." *Econ. Letters* 29(1989):325–31.
- Castle, E., and I. Hoch. "Farm Real Estate Price Components, 1920–78." *Amer. J. Agr. Econ.* 64(1982):8–18.
- DeBondt, W. F. M., and R. H. Thaler. "Anomalies: A Mean-Reverting Walk Down the Street." *J. Econ. Perspectives* 3(1989):189–202.
- Diba, B., and H. Grossman. "Explosive Rational Bubbles in Stock Prices?" *Amer. Econ. Rev.* 78(1988):520–30.
- Dickey, D., and W. Fuller. "Distribution of the Estimators for Autoregressive Time Series with a Unit Root." *J. Amer. Statist. Assoc.* 74(1979):427–31.
- Engle, R., and C. W. J. Granger. "Cointegration and Error-Correction: Representation, Estimation, and Testing." *Econometrica* 55(1987):251–76.
- Falk, B. "A Search for Speculative Bubbles in Farmland Prices." Dep. Econ., Iowa State University, July 1988.
- Featherstone, A. M., and T. G. Baker. "An Examination of Farm Sector Real Asset Dynamics." *Amer. J. Agr. Econ.* 69(1987):532–46.
- Herd, R. W., and W. W. Cochrane. "Farmland Prices and Technological Advance." *J. Farm Econ.* 48(1966):243–63.
- Klinefelter, D. A. "Factors Affecting Farmland Values in Illinois." *IL. Agr. Econ. Res.* 13(1973):27–33.
- Leroy, S., and R. Porter. "The Present Value Relation: Tests Based on Implied Variance Bounds." *Econometrica* 49(1981):555–74.
- Melichar, E. "A Financial Perspective on Agriculture." *Fed. Reserve Bull.* 70(1984):1–13.
- . "Capital Gains versus Current Income in the Farming Sector." *Amer. J. Agr. Econ.* 61(1979):1085–92.
- Phillips, P. C. B., and P. Perron. "Testing for a Unit Root in Time-Series Regressions." *Biometrika* 75(1988):335–46.
- Pope, R. D., R. A. Kramer, R. D. Green, and B. D. Gardner. "An Evaluation of Econometric Models of U.S. Farmland Prices." *West. J. Agr. Econ.* 4(1979):107–19.
- Reynolds, T. E., and J. F. Timmons. *Factors Affecting*

- Farmland Values in the United States*. Iowa State University Agr. Exp. Sta. Res. Bull. No. 566, 1969.
- Said, S., and D. Dickey. "Hypothesis Testing in ARIMA (p, 1, q) Models." *J. Amer. Statist. Assoc.* 80(1985): 369-74.
- Shiller, R. J. "Do Stock Prices Move too Much to Be Justified by Subsequent Changes in Dividends?" *Amer. Econ. Rev.* 71(1981):421-36.
- Sims, C. A. "Macroeconomics and Reality." *Econometrica* 48(1980):1-48.
- Summers, L. H. "Does the Stock Market Rationally Reflect Fundamental Values?" *J. Finance* 41(1986):591-601.
- Tweeten, L. G., and J. E. Martin. "A Methodology for Predicting U.S. Farm Real Estate Price Variation." *J. Farm Econ.* 48(1966):378-93.
- West, K. "Bubbles, Fads, and Stock Price Volatility Tests: A Partial Evaluation." *J. Finance* 43(1988a):639-56.
- . "Dividend Innovations and Stock Price Volatility." *Econometrica* 56(1988b):37-62.

Campaign Contributions and House Voting on Sugar and Dairy Legislation

David G. Abler

In this article, I explore the determinants of campaign contributions by sugar and dairy producer PACs and their role in voting on sugar and dairy legislation in the U.S. House of Representatives. Contributions and voting may be linked in two ways. First, a group can buy votes from a legislator who would otherwise not support the group. Second, it can help elect a legislator who is ideologically predisposed to support the group. Results indicate that the first linkage is not important. Virtually the entire observed association between money and voting is a result of the second linkage.

Key words: farm groups, farm legislation, PACs, rent seeking, voting.

Growing evidence demonstrates the simultaneity between economic conditions and agricultural policy outcomes (Gardner, Lopez, Honma and Hayami). Agricultural policy affects economic conditions in ways that have been exhaustively studied. Economic conditions affect policy because they influence the gains or losses to various political influence groups from a given program and, in turn, the extent to which they lobby for or against that program. In light of the dominant role of policy in agriculture in most countries, an understanding of economic events in agriculture requires an appreciation of both parts of this simultaneous relationship.

Most of the influential farm groups have significant membership in only a small number of districts in the U.S. Congress, and yet they receive sizable political favors. A study of the 1985 farm bill (Abler) found vote trading used heavily by tobacco, sugar, peanut, and dairy farmers, with other farm groups trading votes to a lesser extent. At the same time, many farm groups have political action committees (PACs) that give campaign contributions as part of their rent-

seeking activities. The purpose of this article is to explore the determinants of campaign contributions by sugar and dairy producer PACs and their role in voting on sugar and dairy legislation in the U.S. House.

A group that gives money to a legislator obviously expects something in return. The return can occur in two ways. First, the group can buy votes from a legislator already in office, one who would not support the group without the contribution. Second, it can use contributions to help elect a legislator who, because of personal ideology or some other reason, is predisposed to support the group. If elected, this legislator supports the group regardless of the amount it contributed. The first linkage is considered the most important, but there is little hard evidence one way or another. This article tests if either or both of these linkages is present.

Votes in the House on sugar and dairy amendments to the 1985 farm bill are studied. Of the farm groups, only sugar, dairy, cotton, and cattle producers give campaign contributions to enough representatives to do statistical work in which contributions is a dependent variable. Of these four groups, only sugar and dairy farmers faced amendments specifically directed at them. A group-specific amendment is necessary for a clean test of linkages between contributions and voting. With a motion affecting several groups, the impact of each group on a representative's vote must be distinguished.

David G. Abler is an assistant professor, Department of Agricultural Economics and Rural Sociology, Pennsylvania State University.

The author wishes to thank Cathy Hamlett, Spiro Stefanou, and anonymous *Journal* referees for many helpful suggestions on earlier drafts of this article.

The Model

To analyze the determinants of campaign contributions and voting, consider the following simultaneous-equation vote-contributions model:

$$(1) \quad y_{1t}^* = \gamma_1 E(y_{2t}) + \beta_1' x_{1t} + u_{1t},$$

$$(2) \quad y_{2t} = \gamma_2 E(y_{1t}^*) + \beta_2' x_{2t} + \sigma u_{2t}.$$

Here, y_{1t}^* is a latent variable indicating the propensity of the t th representative to vote for an interest group-specific piece of legislation. The observed counterpart of y_{1t}^* is

$$(3) \quad y_{1t} = \begin{cases} 1, & y_{1t}^* > 0 \\ 0, & y_{1t}^* \leq 0, \end{cases}$$

where $y_{1t} = 1$ for a yes vote; $y_{1t} = 0$ for a no vote; y_{2t} is the net amount of campaign money received by the representative from the interest group, defined as the amount contributed to that representative minus the amount given to his or her electoral opponent(s).

Previous studies generally have treated contributions to electoral opponents as an exogenous variable in an equation explaining gross contributions to the representative or ignored them entirely. Neither approach seems satisfactory. Ideally, the model would have separate equations for contributions to the representative and contributions to his or her opponents. A single net contribution equation is used to keep the model econometrically tractable.¹

A vector of variables, x_{1t} , indicates constituent interests in the motion and the representative's personal ideology. A vector of variables, x_{2t} , measures the legislative power of the representative if elected and the impact of campaign contributions on his or her probability of election. These variables capture the expected returns to a group from a contribution. In addition, γ_1 , γ_2 , and σ are coefficients; β_1 and β_2 are vectors of coefficients; and u_{1t} and u_{2t} are normally distributed random errors with zero means and unit variances.

Equation (1) expresses the propensity to support the legislation as a function of expected net

campaign contributions and exogenous variables. Equation (2) expresses net contributions as a function of expected support for the legislation and exogenous variables. These expected values are obtained by taking expectations of equations (1) and (2), and solving the two-equation system for $E(y_{1t}^*)$ and $E(y_{2t})$.² Expectations are used because neither the representative nor the group can be certain of what the other is doing. The representative does not know what future contributions will be when he or she votes. The group cannot be sure what the representative will do when it contributes prior to the vote.³

Both the legislative motions considered here would have harmed their respective interest groups. If the first linkage between contributions and voting is true (simply buying votes from someone already in office), then $\gamma_1 < 0$. In this case, additional money reduces the likelihood of voting for the motion regardless of constituent interests or the representative's personal preferences. If this linkage is false, $\gamma_1 = 0$. If the second linkage is true (helping to elect someone predisposed to support the group), then $\gamma_2 < 0$. In this case, money is more likely to go to someone whose constituent interests or personal preferences dictate an opposition to the motion. If this linkage is false, $\gamma_2 = 0$.

As a caveat, several interest groups may have a stake in the legislation.⁴ A single group is used here to keep the model clearly focused and econometrically tractable.

Data and Variables

Votes on two amendments to the 1985 farm bill are studied. The first, an anti-sugar motion, would have lowered the loan rate of 18¢ per pound for raw cane sugar in 1985 to 15¢ per pound in 1988 (with the same percentage reduction for beet sugar). It was defeated 147–268 (35%). The second, an anti-dairy motion, would have reduced the milk price support 50¢ per hundredweight per year if government purchases exceeded 5 billion pounds. It would also have eliminated the bill's increases in payment

¹ In this study, counterintuitive results were obtained when gross contributions were used instead of net contributions. Although the impact of voting on contributions was of the expected sign, the impact of contributions on voting was not. The problem may reflect several cases where both a representative and his or her opponent(s) received a lot of money, generating a large value for gross contributions but a small or negative value for net contributions. Many of these representatives voted for the anti-sugar or anti-dairy legislation.

² The solutions are $E(y_{1t}^*) = (\beta_1' x_{1t} + \gamma_1 \beta_2' x_{2t}) / (1 - \gamma_1 \gamma_2)$ and $E(y_{2t}) = (\gamma_2 \beta_1' x_{1t} + \beta_2' x_{2t}) / (1 - \gamma_1 \gamma_2)$. The model's consistency requirement is $\gamma_1 \gamma_2 < 1$.

³ As explained below, both contributions before and after the vote are included in this study. An extended analysis would look at voting and contributions in a dynamic setting.

⁴ Examples are the Sugar Users Group and the Milk Industry Foundation. However, their impact on the 1985 farm bill was definitely limited (Brown, pp. 174–75).

differentials in thirty-three milk marketing order districts. It was defeated 167–247 (40%). Data sources and definitions for all variables are in the appendix.

Contributions during 1983–86 by PACs affiliated with sugar and dairy farmers to representatives of the 99th Congress (1985–86) and their electoral opponents (in 1987 dollars) are used.⁵ Contributions both before and after the votes are used to help avoid prejudging the direction of causality between money and voting. Omitted from the analyses are two representatives serving by 1985 appointment, because they could not have received campaign money in 1984, and the House Speaker. Also omitted are all those who ran for the Senate or state office in 1986, as well as those who retired voluntarily or died between the time of the votes and the 1986 election.

A breakdown of the representatives by voting and net contributions is in tables 1 and 2. Voting and money are clearly associated. The large majority of representatives with positive net contributions voted against the motions. Very few representatives with zero or negative net contributions voted against them. The behavior of those who abstained or voted for these motions in spite of having positive net contributions can be partly understood by looking at table 3. It shows average net contributions for those whose net contributions were nonzero. Those who abstained or voted for the motions received much less on average than those who voted against them. Among those whose net contributions were negative, those who voted no had slightly less spent against them than those who voted yes.

There are too few abstentions to study the determinants of abstaining, so they are excluded from the simultaneous-equation analyses.⁶ Clearly, though, the decision to vote can be as important as the decision about how to vote.

Exogenous variables in the vote equation include the representative's party, his or her "conservatism" (an index from 0 to 1), and the log of one plus group size. Group size is the number of farms in the sugar or dairy standard industrial classification (SIC). District-level data on farm

Table 1. Sugar: Voting and Campaign Contributions

Net Campaign Contributions	Vote			Total
	No	Abstain	Yes	
Positive	224	10	49	283
Zero	12	2	77	91
Negative	6	1	10	17
Total	242	13	136	391

Table 2. Dairy: Voting and Campaign Contributions

Net Campaign Contributions	Vote			Total
	No	Abstain	Yes	
Positive	202	8	76	286
Zero	17	5	60	82
Negative	5	2	16	23
Total	224	15	152	391

Table 3. Average Net Contribution Per Representative

Vote	Net Contributions			
	Sugar		Dairy	
	Negative	Positive	Negative	Positive
No	\$-271	\$2,694	\$-4,224	\$11,642
Abstain	-155*	1,834	-2,074*	5,208
Yes	-320	737	-4,596	3,552

Note: The figures marked with an asterisk should be viewed with caution because they are averages over only 1 or 2 individuals.

organization membership and the number of PAC contributors are not publicly available. Both group size in the representative's district and in his or her state are included. The latter controls for spillover effects from one district to another within a state. Also included is the log of one plus other farms in the district, defined as the total number of farms minus the number in the sugar or dairy SIC. This variable is included to control for vote trading (Abler). Other exogenous variables are the fraction of adults with at least a high school education in the district and the representative's region.⁷

Exogenous variables in the contribution equa-

⁵ Some of the PACs represent more than just farm interests. For example, the American Sugarcane League and Florida Sugarcane League include processors. However, they are predominantly producer organizations (Browne, p. 30).

⁶ A simple option would be to model voting as an ordered probit equation, with an abstention between a yes and a no. However, even this option is precluded by the limited number of abstentions. For sugar, the Hessian was not negative definite at the apparent maximum of the log-likelihood function.

⁷ The log of median family income was also tried. Together, neither income nor education was statistically significant. Separately, education dominated income in terms of asymptotic *t*-ratios.

tion include party, conservatism, the representative's "extremism" (an index from 0 to 1), the log of seniority, and a dummy for incumbency.⁸ Extremism measures deviations from the middle of the road in either direction, so that extreme liberals and extreme conservatives score near 1. More extreme legislators may find the compromises necessary to build successful coalitions less appealing and so may suffer in terms of political influence. Seniority is the reciprocal of the representative's rank within his or her party and is a measure of political power. Incumbency, set equal to 1 if the representative began service prior to November 1984, is a measure of the impact of money on the probability of election. Incumbents have well-known electoral advantages over challengers.

Net contributions are measured as $\delta(m) \log(1 + |m|)$, where m is the net amount of campaign money received and $\delta(\cdot)$ is the sign function. Log transformations here and on the group size variables mitigate against outliers.

Results

Maximum likelihood estimates of the simultaneous-equation systems are in table 4. The results for sugar and dairy tell the same story with respect to the linkages between voting and contributions. The expected propensity to vote for an anti-group motion has a negative and statistically significant effect on group contributions. For each group, money apparently goes to people who like its programs. However, expected net contributions have a negative and statistically insignificant impact on voting. Neither sugar nor dairy PACs seem to be able to buy votes directly from representatives. Virtually the entire association between money and voting appears to operate through the election of those predisposed to support sugar or dairy programs.

These results on the direction of causality between money and voting are consistent with Chappell. Looking at several issues, including dairy price supports in 1975, he found little effect of contributions on voting after accounting for correlated errors in the vote and contribution equations. The results are also consistent with

Peltzman, who found that ideological and party affiliation differences among senators were largely explained by characteristics of campaign contributors and constituents as a whole. Contrast the results with single-equation models, which typically show that money has a large effect on voting (e.g., Welch).⁹

Incumbents start out with large advantages over challengers, as is well known, and the results suggest that they are augmented by sugar and dairy PACs. Interestingly, seniority is not an important factor in contribution decisions. Those with more seniority have more power and can exercise some control over other members, but this may give them a greater ability to ignore PACs. Although extremism has the expected sign, it is not statistically significant.

There is weak evidence that conservatives receive more money, which may result from their political clout during the first half of the 1980s. Party does not seem to influence contribution decisions. Democrats had firm control of the House in 1985, but this may have been offset by the Republicans' greater access to President Reagan. The results for conservatism and party suggest that sugar and dairy contributions can depend on the larger political environment as well as the specifics of a representative and his or her constituents.

In the vote equation, spillover effects from one district to another within a state appear important for sugar but not dairy. Sugar growers received support from other farms, whereas dairy farmers did not. Education is positively associated with support for the anti-sugar and anti-dairy motions. This makes sense because the more educated have higher incomes on average and likely are more knowledgeable about the impacts of sugar and dairy programs.

Conclusions

In this article, I have explored the determinants of campaign contributions by sugar and dairy PACs and their role in voting on sugar and dairy legislation in the U.S. House of Representatives. Voting on anti-sugar and anti-dairy amendments to the 1985 farm bill and campaign contributions in 1983–86 by PACs affiliated with sugar and dairy farmers were studied. A strong

⁸ Membership on the House Agriculture Committee and the sugar or dairy subcommittee were also tried. However, there are too few members on each subcommittee or even the committee to use these variables. For both sugar and dairy, the Hessian was not negative definite at the apparent maximum of the log-likelihood function when either variable was included.

⁹ Single-equation probit estimates here of the impact of net contributions on voting yield coefficients (absolute t -ratios) of -0.22 (7.8) for sugar and -0.11 (6.3) for dairy.

Table 4. Simultaneous Equation Results

Variable	Sugar		Dairy	
	Vote	Contributions	Vote	Contributions
Expected propensity to vote yes		-1.02* (3.7)		-2.15* (3.6)
Expected net contributions	-0.11 (1.1)		-0.004 (0.0)	
Party (Dem. = 0, Rep. = 1)	0.36 (1.0)	-0.52 (0.7)	0.72 (1.4)	-1.19 (1.1)
Conservatism (between 0 and 1)	-0.06 (0.1)	1.39 (1.3)	0.64 (1.3)	2.49 (1.6)
Extremism (between 0 and 1)		-0.75 (1.2)		-1.23 (1.2)
Log (seniority)		-0.17 (1.0)		0.04 (0.2)
Incumbent (No = 0, Yes = 1)		2.48* (4.0)		1.65* (2.1)
Log(1 + group size): In district	-0.07 (0.7)		-0.21* (2.2)	
In state	-0.13* (2.9)		-0.04 (0.6)	
Log(1 + other farms)	-0.13* (4.4)		0.02 (0.5)	
Education (fraction with high school)	5.19* (4.0)		2.49* (2.2)	
Regional dummies:				
Mid-Atlantic	-0.61 (1.3)	0.11 (0.1)	-0.61 (1.2)	1.11 (0.8)
EN Central	-0.79 (1.5)	1.76 (1.8)	-1.13 (1.8)	1.06 (0.7)
WN Central	-2.60* (3.2)	1.57 (1.1)	-2.09* (2.2)	1.18 (0.6)
S Atlantic	-1.03 (1.6)	2.37* (2.1)	-1.01 (1.9)	0.31 (0.2)
ES Central	-0.56 (0.8)	2.99* (2.4)	-1.08 (1.7)	0.57 (0.3)
WS Central	-1.13 (1.9)	0.78 (0.6)	-1.55* (2.9)	-1.58 (0.9)
Mountain	-1.09 (1.4)	3.36* (2.5)	-1.58* (2.2)	0.59 (0.3)
Pacific	-2.17* (3.2)	1.65 (1.4)	-1.05* (2.2)	-0.34 (0.2)
Intercept	-1.37 (1.4)	-0.04 (0.0)	-0.57 (0.5)	3.41 (1.8)
Efron's R^2	0.42	0.36	0.28	0.23
Sample size		378		376

Note: Absolute values of asymptotic t -ratios are in parentheses. An asterisk denotes a t -ratio at least 2 in absolute value. Maximum likelihood estimates (t -ratios) for σ are 3.11 (27.3) for sugar and 4.31 (27.3) for dairy, compared with sample standard deviations for the net contributions variable of 3.88 and 4.92.

association between money and voting was found. However, the results show that sugar and dairy farmers were unable to use campaign contributions to buy votes from representatives who might otherwise vote against them. The association appears to result almost entirely from the role contributions play in helping to elect people predisposed to support sugar and dairy programs.

The major policy implication to come out of

this study concerns the impact of a limit or ban on PAC contributions, which many have proposed. Given the inability of sugar and dairy PACs to buy votes directly, the results imply that the short-run impact on sugar and dairy policy would be quite small. In the long run, however, representatives not so supportive of sugar and dairy programs could replace current ones as PAC money ceased to be such an important

factor in election campaigns. Thus, the long-run impact could be substantial.

[Received September 1989; final revision received April 1990.]

References

- Abler, David G. "Vote Trading on Farm Legislation in the U.S. House." *Amer. J. Agr. Econ.* 71(1989):583-91.
- Browne, William P. *Private Interests, Public Policy, and American Agriculture*. Lawrence: University Press of Kansas, 1988.
- Chappell, Henry W., Jr. "Campaign Contributions and Congressional Voting: A Simultaneous Probit-Tobit Model." *Rev. Econ. and Statist.* 64(1982):77-83.
- Congressional Quarterly, Inc. *Congressional Quarterly Almanac*, 1984-86. Washington DC: Congressional Quarterly Press, 1985-87.
- Gardner, Bruce L. "Causes of U.S. Farm Commodity Programs." *J. Polit. Econ.* 95(1987):290-310.
- Honma, Masayoshi, and Yujiro Hayami. "The Determinants of Agricultural Protection Levels: An Econometric Analysis." *The Political Economy of Agricultural Protection: East Asia in International Perspective*, ed. Kym Anderson and Yujiro Hayami. Sydney: Allen and Unwin, 1986.
- Lopez, Rigoberto A. "Political Economy of U.S. Sugar Policies." *Amer. J. Agr. Econ.* 71(1989):20-31.
- Peltzman, Sam. "Constituent Interest and Congressional Voting." *J. Law and Econ.* 27(1984):181-210.
- U.S. Department of Commerce, Bureau of the Census. *Congressional District Atlas*, 99th Congress. Washington DC, 1985.
- . *1980 Census of Population and Housing: Congressional Districts of the 98th Congress*. Washington DC, 1983.
- . *1980 Census of Population and Housing: Congressional Districts of the 99th Congress*. Washington DC, 1985.
- . *1987 Census of Agriculture*, vol. I, pts. 1-51. Washington DC, 1989.
- U.S. Federal Election Commission. "Committee Index of Candidates Supported/Opposed," 1983-84 and 1985-86, unpublished data, 1989.
- Welch, W. P. "Campaign Contributions and Legislative Voting: Milk Money and Dairy Price Supports." *West. Polit. Quart.* 35(1982):478-95.
- and 290 (dairy). Source: *Congressional Quarterly Almanac*, 1985.
- Contributions.** Contributions to a representative of the 99th Congress made during 1983-86 minus those made to his or her electoral opponent(s) (in 1987 dollars). They are converted to 1987 dollars with the gross national product (GNP) implicit price deflator. For sugar, nine PACs are covered: American Sugar Beet Growers Association; American Sugarcane League; California Beet Growers Association; Florida Sugarcane League; Great Lakes Sugar Beet Growers Association; Hawaiian Sugar Planters Association; Southern Minnesota Beet Sugar Cooperative; Texas Sugar Beet Growers Association; and U.S. Beet Sugar Association. For dairy, seventeen PACs are covered: Associated Milk Producers Inc.; Dairymen's Ltd. Agricultural Association; Dairymen, Inc.; Georgia, Kentucky, Louisiana, Mississippi, North Carolina, Pennsylvania, Tennessee, and Virginia chapters of Dairymen, Inc.; League of California Milk Producers; Michigan Milk Producers Association; Mid-American Dairymen, Inc.; North Pacific Dairymen's Association; United Dairymen of Arizona; and Western Dairymen's Association. Source: Federal Election Commission, "Committee Index of Candidates Supported/Opposed," 1983-84 and 1985-86.
- Party.** Democrat = 0, Republican = 1. Source: *Congressional Quarterly Almanac*, 1985.
- Conservatism.** Defined as YES/(YES + NO). YES (NO) is the number of conservative coalition roll calls on which the representative cast a roll call vote in agreement (disagreement) with the position of the conservative coalition in the 99th Congress. The conservative coalition is a voting alliance of Republicans and Southern Democrats against Northern Democrats. A conservative coalition roll call is any roll call on which the majority of voting Republicans and the majority of voting Southern Democrats cast a roll call vote opposite to that cast by the majority of voting Northern Democrats. Source: *Congressional Quarterly Almanac*, 1985 and 1986.
- Extremism.** Defined as $2|C - 0.5|$, where C is the value of the conservatism index above.
- Seniority.** The reciprocal of the representative's rank within his or her party. Source: *Congressional Quarterly Almanac*, 1985.
- Incumbent.** Equals 0 if the representative began service in November 1934 or later; equals 1 otherwise. Source: *Congressional Quarterly Almanac*, 1985.
- Farm Groups.** The number of farms in an appropriate standard industrial classification (SIC), 1987. The SICs used are sugar (0133) and dairy (024). Congressional district data on farm numbers were constructed from county data in the 1987 Census of Agriculture. For sugar, where county data are unavailable, the following approximation is used: $SIC_{ij} = (TOTAL_{ij}/TOTAL_j)SIC_j$, where i is state and j is county. $TOTAL$ is the total number of farms producing the commodity, while SIC is the number in the commodity's SIC. The number of other farms is the total number of farms of all types minus the number in the sugar or dairy SIC.

Appendix

Data Sources and Variable Definitions

Vote. Each of the following actions is counted: a roll call vote; a specific (but not general) paired vote; a publicly announced position; and a response to a Congressional Quarterly poll. The CQ House vote numbers are 289 (sugar)

The county data were aggregated to congressional district data using maps of congressional districts and tables listing the congressional districts(s) in each county in the U.S. from the *Congressional District Atlas*, 99th Congress. For a county with more than one district in it, the totals for that county were apportioned among districts in accordance with a rough guess of the percentage of the county's rural area in each district.

Education. Fraction of persons aged twenty-five and over who have completed high school, 1980. sources: *1980 Census of Population and Housing: Congressional Districts of the 98th Congress*; and *1980 Census of Population and Housing: Congressional Districts of the 99th Congress*.

Regional Dummies. The Census Bureau regions are used. The control region is New England.

The Aggregate Effects of Risk in the Agricultural Sector

Jack Meyer and Lindon J. Robison

The theory of the firm facing a random output price is extended to include industry equilibrium conditions in a way particularly important for agriculture. Land prices adjust to maintain industry equilibrium. The comparative statics and the effects of policy for the firm are significantly altered. The main finding is that risk and other parameter changes are capitalized into the price of land and yield wealth or income effects rather than substitution effects. This often simplifies the analysis.

Key words: competitive firm, land prices, mean-variance, risk.

Empirical research has demonstrated that adjustments in land prices are an important equilibrating mechanism for the agricultural sector. Public policies and farm programs are known to affect land prices (e.g., Vantresse, Reed, and Skees). Changes in the riskiness or level of returns to agricultural production are also reflected in land prices (Robison, Lins, and VenKataraman). Developing a single aggregate model under risk which displays the linkages between risk, return, and land prices is our main task in this paper. The purpose is to understand better the effects of public policies which alter the returns and risk faced by the agricultural producer.

This study also adds to the vast literature concerning a competitive firm facing a random output price. Equilibrium conditions are obtained for such a firm operating in a competitive industry where an input's price adjusts to maintain industry equilibrium. The findings complement the recent work of Appelbaum and Katz, who analyze competitive firms facing output price risk when operating in a constant cost industry.

The competitive agricultural firm modeled here uses land and other inputs in producing a product whose price is random. The price of land is assumed to rise or fall to maintain equilibrium in the industry. Two approaches, an asset pricing model and explicit supply and demand re-

lationships for land, are used to determine equilibrium land prices.

The comparative statics for the individual firm in this model are quite different from those obtained by Sandmo and others who ignore industry equilibrium considerations. This paper's results also differ from those of Appelbaum and Katz, who assume a different industry equilibrating mechanism.

The paper is organized as follows. First, the literature concerning the competitive firm in a competitive industry is reviewed, and a simple model of a competitive firm facing a random output price and using land as an input is presented. The next section develops conditions for equilibrium in a competitive industry when land is in fixed supply and its price adjusts to maintain equilibrium in the industry. Then, the model is expanded to reflect more adequately the conditions in the agricultural sector. Finally, a general policy example highlights the differences between comparative static results in this model and those obtained by others.

The Competitive Firm in a Competitive Industry

A large literature concerning the competitive firm operating in a competitive industry exists. The starting point is a model where the firm chooses output level x to maximize profit, π , where $\pi = p \cdot x - c(x)$. In this model, p represents the output price and $c(x)$ is the cost function.

Without randomness, the industry is assumed to be in equilibrium when the firm earns zero

Jack Meyer and Lindon J. Robison are professors of economics and agricultural economics, respectively, at Michigan State University.

Funding was provided jointly by the Michigan Agricultural Experiment Station and Economic Research Service, U.S. Department of Agriculture under research agreement number 58-3J23-4-00276.

profit. For the constant cost industry case, $c(x)$ is exogenous to both the firm and the industry and only output price can adjust to attain zero profit. The standard assumptions are that p depends on the industry output level Q and that entry or exit occurs until Q implies zero profit for each firm, i.e., $p(Q) = c(x)/x$.

Sandmo extends the competitive firm model to the case where output price is a random variable and expected utility of profit is maximized. The effects of shifting various parameters, including the mean and the riskiness of the output price, are also determined. Comparative statics in this model differ from the certainty case; moreover, these differences have received much attention in the agricultural economics as well as the economics literature.

Recently, Appelbaum and Katz note that Sandmo fails to consider equilibrium in the competitive industry. That is, how the individual firm reacts to changes in certain parameters is calculated without determining whether or not those parameter changes are consistent with equilibrium in the competitive industry.

Appelbaum and Katz add to Sandmo's analysis by including an output price which is a random variable whose mean depends on industry output level. They assume that output price is given by $p(Q) + \epsilon$, where ϵ is random with a zero mean, and $p(Q)$ is a nonrandom term giving the mean output price as a decreasing function of industry output level. Costs are assumed exogenous to the firm and to the industry.

Industry equilibrium is attained when no firm desires to enter or exit the industry. Because profit is a random variable, it cannot be set equal to zero. Instead, Appelbaum and Katz assume that firms can attain a reservation level of expected utility in other activities and choose to enter or leave this industry if the expected utility from profit exceeds or falls short of this reservation level.

This industry equilibrium requirement implies that the mean output price depends on all other parameters in the model, a link which does not exist in the Sandmo formulation. Thus, shifts in those parameters have effects in addition to those determined by Sandmo. Consequently, Appelbaum and Katz find the comparative statics for the firm are significantly altered when industry equilibrium conditions are considered.

For certain industries a second mechanism for adjusting to industry equilibrium exists. Adjustments in input prices can alter a firm's cost function, thus altering the firm's desire to enter

or exit the industry.¹ This mechanism is particularly relevant in agriculture because adjustments in land prices, and hence costs of production, alter the profitability of firms in the industry.

When microeconomists wish to focus exclusively on this second mechanism for attaining industry equilibrium, a simple model with only one input in fixed supply is often considered. Furthermore, only the price of this input is assumed to adjust to induce entry or exit from the industry.

Formally, in this one input model, profit is given by $\pi = p \cdot L - \phi \cdot M \cdot L$, where L is the input level chosen by the firm. Constant returns to scale is assumed and one unit of input produces one unit of output.² As is usual under the constant cost assumption, firm size is indeterminate. Because land is a durable, its cost is written as a per period cost $\phi \cdot M$, where ϕ is the risk-free interest rate and M is the durable's price.

The competitive firm treats input and output prices as exogenous variables. For the industry however, the input price adjusts to maintain industry equilibrium. Because input L is available in fixed supply, its price serves as a rationing mechanism. Without randomness, industry equilibrium is again characterized by the zero profit condition. This implies the standard capitalization formula $M = p/\phi$ as the equilibrium price for the durable input. The return per period for the durable is capitalized using the risk-free discount rate.

This model of a competitive firm can be extended to the random output price case. Profit is given by $\pi = p \cdot L - \phi \cdot M \cdot L$ as before, but now p is a random variable. The firm is assumed to choose L to maximize expected utility of profit. If strict risk aversion is assumed, even with constant returns to scale, a finite solution for L obtains and firm size is determinate.

Output price p is the only source of randomness in this model, and profit is a positive linear transformation of p . Thus, all potential profit distributions are location and scale transformations of the distribution of output price. This

¹ When input prices increase to reduce profits earned by firms in an industry, this is often referred to as the increasing cost industry case.

² The case of nonconstant returns to scale can also be considered. The profit expression becomes $\pi = p \cdot f(L) - \phi \cdot M \cdot L$, where $f(L)$ is the production function. $f(L)$ must be such that firms are neither infinitely small nor large, and the zero profit condition requires that $M = p \cdot f(L)/\phi \cdot L$.

property allows the optimal choice of L and much of the comparative static analysis to be conducted in a mean-standard deviation (MS) framework without violating the expected utility maximization assumption or imposing other special assumptions.³

The MS formulation of the decision model with $\pi = p \cdot L - \phi \cdot M \cdot L$ has the competitive firm choosing L to maximize $V(\sigma, \mu)$, where mean $\mu = \mu_p \cdot L - \phi \cdot M \cdot L$ and standard deviation $\sigma = \sigma_p \cdot L$. The notation μ_p and σ_p represent the mean and standard deviation of output price, respectively. Solving for L in the expression for σ and substituting into the expression for μ reduces the pair of constraints to the linear restriction given as equation (1). The firm's choice of L is a selection of some point on this straight line opportunity set in (σ, μ) space.

$$(1) \quad \mu = (\mu_p - \phi \cdot M) \cdot \sigma / \sigma_p.$$

Notice that the firm's opportunities depend on the four parameters, μ_p , σ_p , ϕ , and M .

As in the model without randomness, industry equilibrium considerations imply that the price of the input L adjusts to attain the appropriate profit level. Because profit is random, the equilibrium land price is more difficult to define. The next section first uses an asset-pricing model to resolve the issue and then develop an alternate approach involving supply and demand equations for land.

Input Price Adjustment and Industry Equilibrium

The input L is an asset which generates a random rate of return p/M in each time period. The level of the asset price, M , which implies an equilibrium random rate of return, must be resolved. The random return to L is similar to the return from holding one unit of a financial asset, such as a share of stock. The capital asset pricing model (CAPM) suggests that the price of the asset or input adjusts until its mean rate of return is equal to the risk-free rate of return plus a risk adjustment factor. This factor depends on the market compensation for risk, $[(\mu_m - \phi)/\sigma_m]$, the diversification characteristics of the asset, ρ , and the riskiness of the return, (σ_p/M) . In these expressions, ϕ is the risk-free rate of return, μ_m

and σ_m are the mean and standard deviation of the rate of return on the market portfolio, and ρ is the correlation coefficient between the rate of return on this asset and the rate of return on the market portfolio.

Formally applying the CAPM to the pricing of the input L requires that the mean return per dollar satisfies⁴

$$(2) \quad \begin{aligned} \mu_p/M &= \phi + [(\mu_m - \phi)/\sigma_m] \\ &\cdot \rho \cdot (\sigma_p/M) = \phi + \eta \cdot (\sigma_p/M). \end{aligned}$$

For notational convenience, η represents both the market compensation for risk and the asset's diversification possibilities.⁵ The η term will be treated as fixed throughout this analysis.

The CAPM assumes that the asset is part of a maximally diversified portfolio called the market portfolio. Levy has demonstrated that a pricing formula exactly like (2) holds even with a limited ability to diversify. The coefficient η still represents both the market price of risk and the diversification possibilities of the asset. Now, however, the $[(\mu_m - \phi)/\sigma_m]$ and ρ terms are calculated with respect to a less well-diversified portfolio than the market portfolio. It makes no difference in this analysis whether (2) results from CAPM or Levy's argument. Thus, the interpretation of η is left purposely ambiguous.⁶

The equilibrium value for M which pricing equation (2) implies is given by

$$(3) \quad M = (\mu_p - \eta \cdot \sigma_p) / \phi = \mu_p / \phi - \eta \cdot \sigma_p / \phi.$$

As in the example without risk, the industry equilibrium condition provides an additional link among the parameters in the model and quite different comparative statics result.

Before exploring the implications of this additional equilibrium condition, an alternate approach not using CAPM is presented. This second method of modeling the equilibrium price for land specifies a particular functional form for each firm's preferences in order to calculate an explicit demand function for L . The linear mean-variance utility function is assumed. That is, the i th firm is assumed to choose L_i to maximize $\mu_i - (\lambda_i/2)\sigma_i^2$, where $\mu_i = \mu_p \cdot L_i - \phi \cdot M \cdot L_i$ and $\sigma_i = \sigma_p \cdot L_i$. The parameter λ_i represents the risk aversion measure of the firm and

⁴ Since one unit of output is produced per unit of input, μ_p and σ_p are also the mean and standard deviation of return per unit of input.

⁵ In standard CAPM notation, the η term is not equal to the traditional CAPM beta (β) but is related to it in the following way: $\eta = (\mu_m - \phi) \cdot \beta \cdot \sigma_p / \sigma_p$.

⁶ Barry has treated the pricing of agricultural land using both CAPM and Levy's pricing argument.

³ Sinn and Meyer describe the implications of this location and scale property. Meyer and Robison use it to analyze the hedging behavior of the competitive firm.

is assumed positive. Choosing L to maximize this function leads to

$$(4) \quad (\mu_p - \phi \cdot M) - \lambda_i \cdot \sigma_p^2 \cdot L_i = 0$$

as the first-order condition, with the second-order condition satisfied since λ_i is positive.

Solving for the optimal choice of L_i , one obtains

$$(5) \quad L_i = (\mu_p - \phi \cdot M) / (\lambda_i \cdot \sigma_p^2).$$

Thus, the individual producer's demand for L depends on all of the parameters in the model: the mean and variance of output price, the risk-free rate of return, the input price, and the producer's risk aversion level. Only this latter term is firm specific.

The demand functions for the durable input can be easily added across n firms. Doing so, one obtains

$$(6) \quad L_D = ((\mu_p - \phi \cdot M) / \sigma_p^2) (1/\lambda_1 + \dots + 1/\lambda_n)$$

as the aggregate demand function for the n producers.

To find the equilibrium price for L , this demand function is set equal to the given supply L_o . Solving for M one obtains

$$(7) \quad M = (\mu_p / \phi) - (L_o \cdot \sigma_p^2) / (\phi \cdot (1/\lambda_1 + \dots + 1/\lambda_n)).$$

The form of this pricing equation is similar to equation (3) based on the CAPM argument. That is, in equilibrium, the price of L equals a capitalization of the mean return, plus a term which discounts for the riskiness of the production activity and the diversification possibilities of spreading the risk across the n producers. For this supply and demand case, the equivalent of η is the term $(L_o \cdot \sigma_p^2) / (1/\lambda_1 + \dots + 1/\lambda_n)$. This expression represents the diversification possibilities of distributing $L_o \cdot \sigma_p^2$ units of risk across n different individuals with risk aversion measures λ_i . As the agricultural producer's risk aversion goes to zero, the premium caused by the riskiness of profits also goes to zero.

338.105
Am 35

Firm Behavior Constrained by Industry Equilibrium

Comparative static analysis in the model of the firm is significantly altered when industry equilibrium is considered. Recall that μ_p , σ_p , ϕ , and M together define the firm's opportunities in

equation (1) and, hence, affect its choices. Under industry equilibrium, these four parameters are linked by an additional equation, either (3) or (7), depending on whether the asset-pricing or supply and demand argument is employed. Interestingly, this additional equation actually simplifies the analysis.

For instance, when the value for M given by (3) is substituted into (1), the constraint the firm faces becomes $\mu = \eta \cdot \sigma$. Thus, only the market price of risk and the diversification parameter affect the firm's opportunities. Changing the mean or riskiness of output price is offset by the resulting adjustments in M so that the firm's opportunity set remains unchanged.

Similar simplification occurs when the equilibrium value for M arises from supply and demand analysis. In this case, (7) is substituted into the first-order condition for the firm, (4). This yields $L_o / (1/\lambda_1 + \dots + 1/\lambda_n) - \lambda_i \cdot L_i = 0$. The firm's choice of L_i no longer depends on the mean or variance of output price but only on the available supply of land and the firm's risk-aversion level relative to others in the industry.

This model of the competitive firm is adequate to discuss the impact of industry equilibrium on the price of land. It is not, however, sufficiently well developed to consider the agricultural policy implications of this aggregate risk model. Thus, the model is augmented with several other parameters to enrich the analysis.

First, the initial wealth of the firm is assumed nonzero and held as cash and land. Thus, parameter changes affect the agricultural firm's wealth through their impact on land prices.

Let $C_o + M \cdot \bar{L}$ be the firm's initial wealth, where C_o represents the cash portion, and $M \cdot \bar{L}$ is the land wealth. Producers own different amounts of land but can use quantities of land which differ from the amount owned. With this change one can distinguish between the effects of various programs on producers who rent versus those who own their land. Because nearly three-fourths of the agricultural sector's assets consists of land, this change in the model is particularly relevant.

The second change in the model is to introduce other inputs in the production of output. Even if the land quantity is fixed, industry output can expand or contract by altering other inputs. Formally, output level is assumed to be given by $L \cdot f(s)$, where s represents another input or a vector of inputs. These inputs are measured per acre of land. As before, this production function represents a constant returns-to-scale

P 7688

process with output per acre depending on s .⁷ The function $f(s)$ is increasing and concave in its arguments.

The producer's objective is to choose L and s to maximize expected utility from terminal wealth W given by

$$(8) \quad W = (1 + \phi)(C_o + M \cdot \tilde{L}) + (p \cdot f(s) - \phi \cdot M - p_s \cdot s) \tilde{L},$$

where p_s is the vector of prices for inputs s . The producer takes all prices as parameters outside its control. This wealth function is linear in the random variable, and hence MS analysis is appropriate.

In the MS framework, the producer chooses L to maximize $V(\sigma, \mu)$, where $\mu = (1 + \phi)(C_o + M \cdot \tilde{L}) + (\mu_p \cdot f(s) - \phi \cdot M - p_s \cdot s)L$ and $\sigma = \sigma_p \cdot f(s) \cdot L$. Again, one can solve for L in the second restriction and substitute into the first to define a linear opportunity set in (σ, μ) space.

$$(9) \quad \mu = (1 + \phi)(C_o + M \cdot \tilde{L}) + (\mu_p \cdot f(s) - \phi \cdot M - p_s \cdot s)\sigma/(\sigma_p \cdot f(s)).$$

All model parameters affect the producer's opportunity set and many impact the slope as well as the intercept. Equation (9) is the producer's constraint if industry equilibrium considerations are ignored. Under risk aversion, $V(\sigma, \mu)$ is known to be quasi-concave and the second-order conditions for the maximization are satisfied.

The level of s only affects the slope of (9). Thus, all risk-averse producers choose s to maximize this slope independent of their risk attitudes. On the other hand, producers choose different levels of L depending on their risk-taking characteristics. Each value for L is represented by a particular point on the linear opportunity set in (σ, μ) space.

The level of s which maximizes the slope of (9) is defined by $(\phi \cdot M + p_s \cdot s)/f(s) = p_s/f'(s)$. The left and right sides of this equality are the producer's average cost and marginal cost per acre, respectively. Thus, all producers choose to operate where the average cost of production is at its minimum. However, changes in various parameters can affect this level of output per acre. Specifically, changes which induce increases in the price of land will cause the output per acre to rise.

As in the simple model, other input prices and output prices are assumed exogenous to the in-

dustry. Only the price of land adjusts to maintain industry equilibrium. The resulting equilibrium price for land is given by an equation similar to (3) or (7) found earlier. The CAPM pricing argument applied to this more general model requires that the rate of return to land satisfy

$$(10) \quad (\mu_p \cdot f(s) - p_s \cdot s)/M = \phi + \eta \cdot \sigma_p \cdot f(s)/M.$$

The rate of return to land is the value of output per acre exclusive of other input costs divided by the price of land. The mean, standard deviation and correlation are calculated for this term. Solving equation (10) for M leads to a land-pricing equation.

$$(11) \quad M = (\mu_p \cdot f(s) - p_s \cdot s - \eta \cdot \sigma_p \cdot f(s))/\phi.$$

Similarly if one assumes the explicit linear mean-variance preference function and requires that the demand for land is equal to a fixed supply, the price of land satisfies

$$(12) \quad M = ((\mu_p \cdot f(s) - p_s \cdot s)/\phi) - (L_o \cdot (\sigma_p \cdot f(s))^2)/(\phi \cdot (1/\lambda_1 + \dots + 1/\lambda_n)).$$

To capture the industry equilibrium condition arising under the CAPM argument, land-pricing equation (11) is substituted into the firm's linear constraint (9). The opportunity set for the firm then is

$$(13) \quad \mu = (1 + \phi)(C_o + (\mu_p \cdot f(s) - p_s \cdot s - \eta \cdot \sigma_p \cdot f(s)) \cdot \tilde{L}/\phi) + \eta \cdot \sigma.$$

Changes in all parameters other than η only shift the intercept of this linear constraint, their effect is equivalent to a change in initial wealth. For this simple model, when agriculture is in competitive equilibrium, the effects of changing the mean or riskiness of output price are capitalized into the price of land. Parameter changes have no effect if $\tilde{L} = 0$, i.e., if the producer is not a land owner. Hence, the impact of these changes depends on the producer's land ownership position.

Equation (13) indicates that shifting any parameter except τ is represented by a parallel shift in the opportunity set in (σ, μ) space. The effect of a parallel shift in the opportunity set depends on the producer's risk-aversion characteristics (Meyer and Robison). If the producer is constant absolute risk averse, no change in the optimal L occurs, while acreage increases (declines) under decreasing (increasing) absolute risk aversion.

If the supply and demand model for land is

⁷ If constant returns to scale is not assumed, the analysis becomes considerably more complicated.

used to determine an equilibrium price for land, the analysis is less interesting because wealth effects in the linear mean variance model do not affect the producer's choice of L when λ_i is treated as a constant. It is possible, however, to treat λ_i as a variable in order to analyze this case more fully (Robison and Barry). The optimal level of s and L are defined by the equations:

$$\mu_p \cdot f'(s) - p_s - \lambda_i \cdot \sigma_p^2 \cdot f(s) \cdot f'(s) \cdot L_i = 0 \text{ and} \\ L_o / (1/\lambda_1 + \dots + 1/\lambda_n) - \lambda_i \cdot L_i = 0.$$

A Policy Application

The federal government provides many opportunities to alter the mean and/or riskiness of output price through participation in government-sponsored programs (Gardner, Gardner et al., Fleisher). A general example of these programs is used here to illustrate how policy evaluation could change when industry equilibrium is considered.

Suppose government intervention in commodity markets attempts to maintain or increase the mean price while reducing its variability and that program participants are required to idle $(1 - \alpha)$ percent of the land. That is, $\mu_p = \mu_p(G)$ and $\sigma_p = \sigma_p(G)$, where G is an exogenously determined level of government intervention. If the idled land produces nothing of value, then the opportunity cost to the producer of using land increases, and is $\phi \cdot M/\alpha$ rather than $\phi \cdot M$.

When these changes are introduced into (9), equation (14) results. This is the product's constraint ignoring industry equilibrium conditions. The G and α parameters alter the slope but not the intercept of this line.

$$(14) \quad \mu = (1 + \phi)(C_o + M \cdot \bar{L}) + (\mu_p(G) \cdot f(s) - \phi \cdot M/\alpha - p_s \cdot s) \sigma / (\sigma_p(G) \cdot f(s)).$$

Analysis of the firm's response to changes in G or α involve income and substitution effects, and the results are often ambiguous. Also, because \bar{L} only enters the intercept term, the impact of changes in G or α do not depend on the extent to which producers own their land.

When industry equilibrium is considered, the price for land must be adjusted to incorporate these government effects. The CAPM argument requires that the net return per unit of land satisfy

$$(15) \quad \alpha \cdot (\mu_p(G) \cdot f(s) - p_s \cdot s) / M \\ = \phi + \eta \cdot \alpha \cdot \sigma_p(G) \cdot f(s) / M,$$

where the return per acre is adjusted to account for the $1 - \alpha$ percent which is idled. Solving for M yields

$$(16) \quad M = \alpha \cdot (\mu_p(G) \cdot f(s) - p_s \cdot s - \eta \cdot \sigma_p(G) \cdot f(s)) / \phi.$$

When this value for M is substituted into the producer's constraint, (14), the resulting opportunity set displays far different sensitivities to government action. The new opportunity set is

$$(17) \quad \mu = (1 + \phi)(C_o + \alpha \cdot (\mu_p(G) \cdot f(s) - p_s \cdot s - \eta \cdot \sigma_p(G) \cdot f(s)) \cdot \bar{L} / \phi) + \eta \cdot \sigma.$$

Now the government's actions only shift the intercept and do not affect the slope of the linear constraint. The effect of these changes can generally be determined on the basis of an assumption concerning the producer's absolute risk-aversion measure.

The effect of the government's choice of G or α is proportional to the value for \bar{L} . The opportunity set and welfare of producers who only rent land, $\bar{L} = 0$, is not affected by the government's choice of G and α . These producers do make different decisions concerning the level of other inputs since the optimal s depends on M , but their overall opportunities to earn profits are not affected by the government actions. For instance, if μ_p is increased through a support price, land prices rise so that renters find the random rate of return to agricultural production no different than before.

When the producer owns some or all of the land which is used, $\bar{L} > 0$, changes in G or α alter the producer's opportunity set and welfare by changing initial wealth. Changes which increase (decrease) M make the producer better (worse) off. These welfare changes, however, are the result of the gains to these individuals as land owners and not from their agricultural production activities.

Conclusions

Of course this model abstracts from the realities of the situation. Prices other than land prices are likely to adjust in order to maintain industry equilibrium. Nonetheless, this paper extends the theory of the firm facing a random output price to include industry equilibrium conditions in a way particularly important for agriculture. It complements the work of Appelbaum and Katz, who analyze the constant cost industry.

The comparative statics for the firm are significantly altered when industry equilibrium is

considered. For the agricultural firm, the effects of policies which alter the mean or riskiness of output price are different than when industry equilibrium conditions are ignored. The main finding is that risk and other parameter changes are capitalized into the price of land and yield wealth or income effects rather than substitution effects. This greatly alters and often simplifies the comparative statics in the model of the competitive firm.

[Received August 1989; final revision received April 1990.]

References

- Appelbaum, E., and E. Katz. "Measures of Risk Aversion and Comparative Statics of Industry Equilibrium." *Amer. Econ. Rev.* 76(1986):524-29.
- Barry, P. J. "Capital Asset Pricing and Farm Real Estate." *Amer. J. Agr. Econ.* 62(1980):549-53.
- Fleisher, Beverly. *Agricultural Risk Management*. Boulder CO: Lynne Rienner Publishers, 1989.
- Gardner, B. L. *The Governing of Agriculture*. Lawrence: Regents Press of Kansas, 1981.
- Gardner, B. L., R. E. Just, R. A. Kramer, and R. D. Pope. "Agricultural Policy and Risk." *Risk Management in Agriculture*, ed. P. J. Barry, chap. 16. Ames: Iowa State University Press, 1984.
- Levy, H. "Equilibrium in an Imperfect Market." *Amer. Econ. Rev.* 63(1978):643-58.
- Meyer, J. "Two-Moment Decision Models and Expected Utility Maximization." *Amer. Econ. Rev.* 77(1987):421-30.
- Meyer, J., and L. J. Robison. "Hedging Under Output Price Randomness." *Amer. J. Agr. Econ.* 70(1988):268-72.
- Robison, L. J., and P. J. Barry. *The Competitive Firm's Response to Risk*, chap. 18. New York: Macmillan Co., 1987.
- Robison, L. J., D. A. Lins, and R. Venkataraman. "Cash Rents and Land Values in U.S. Agriculture." *Amer. J. Agr. Econ.* 67(1985):794-805.
- Sandmo, A. "On the Theory of the Competitive Firm Under Price Uncertainty." *Amer. Econ. Rev.* 61(1971):65-73.
- Sinn, Hans-Werner. *Economic Decisions Under Uncertainty*. Amsterdam, New York, and Oxford: North-Holland Publishing Co., 1983.
- Vantresse, V., M. Reed, and J. Skees. "Mandatory Production Controls and Asset Values: A Case Study of Burley Tobacco Quotas." *Amer. J. Agr. Econ.* 71(1989):319-25.

Application of Computer Graphics to Undergraduate Instruction in Agricultural Economics

David L. Debertin and Larry D. Jones

This article outlines are experience in building a freshman-level course in agricultural economics employing computer graphics imaging. Lecture material is displayed with a computer connected to a large-screen projector producing high-resolution graphics. The complete course consists of approximately 1,200 computer-generated text, chart, or graphics images. An evaluation of the new method was conducted. Results indicate that most students prefer lectures that employ computer graphics to those that use a chalkboard or an overhead projector. Evidence supports the hypothesis that students perform better on exams when the innovations described in this paper are adopted.

Key words: computer graphics, teaching technology, undergraduate instruction.

It is now feasible to develop and use computer graphics and imaging technologies to teach an entire undergraduate course in agricultural economics. Computer graphics has been used in the past for data analysis (Bay and Schoney; Debertin, Pagoulatos, and Bradford), for the analysis of functions used in agricultural economics research (Debertin, Pagoulatos, and Bradford), and to supplement textbook presentations (Debertin 1985 and 1986). Harris employed computer graphics as an instructional aid in teaching a futures trading course. Computers, sometimes connected to television monitors, have been used in undergraduate agricultural economics courses, particularly in portions of courses requiring spreadsheets, linear programming models, simulators, or other computerized decision aids (Bentley, Hudson et al., Litzenberg, Osburn et al.).

Agricultural economists are convinced of the need for (a) frequent updating of undergraduate curricula (Coffey, Erven, Dobson, Kropp 1973a);

(b) competency-based curricula development (Beck et al., Mather et al.); (c) incorporating computer skills into revamped curricula (Manderscheid, p. 744); and (d) new technologies using microcomputers to improve educational delivery (Harris, Hudson et al., Litzenberg). Discussion in our journals suggests that new teaching technologies, particularly those that use microcomputers, have had greater impacts on adult than on undergraduate education (Brown, Diesslin, Holt, Hughes). Fuller argues (p. 980) that "we can and should be using computers . . . as an integral part of our educational program." Hudson et al. (p. 177) suggest that "continuing advances in technology offer a growing variety of educational uses of the microcomputer not previously considered."

Extension educators, not undergraduate teachers, frequently have been the earliest to adopt the new information-delivery technologies employing computers. Recently, Wetzstein (p. 63) even argued that the new technologies for teaching basic economics, such as television and computers in instruction, "stifle students' imaginations, contribute to a dependent learning style, and fail to stimulate interest in the subject matter." Our experiences with new teaching technologies employing computers and color projection monitors suggest otherwise; these new technologies have helped students better learn principles of basic economics applied to agriculture than when conventional teaching methods were employed.

David L. Debertin is a professor of agricultural economics, and Larry D. Jones is a professor and chair, Department of Agricultural Economics, University of Kentucky.

The investigation reported in this paper, no. 90-1-106, is in connection with a project of the Kentucky Agricultural Experiment Station and is published with the approval of the director.

Debertin was responsible for developing all of the computer graphics images comprising the course and teaches the course in the fall semester; Jones teaches the course in the spring semester, using the computer graphics images and study guide.

Appreciation is extended to Joe T. Davis for helpful comments on an earlier draft of this manuscript.

This paper outlines experiences at the University of Kentucky in building a freshman-level course in agricultural economics employing microcomputer graphics imaging as the basis for every lecture and using the method to teach a class of undergraduate students. As part of the new method, other teaching innovations were also developed, and these are also described in the discussion that follows. Evaluations of student reactions to these innovations and their effect on student performance on examinations were conducted, and the results are presented.

The Methods

All of the notes, charts and diagrams used in the course were developed using computer graphics software.¹ Throughout the semester, lecture material was displayed to students using a computer connected to a large-screen projector capable of producing text and high-resolution (EGA/VGA) graphics. The complete course consists of approximately 1,200 computer-generated text, chart, or graphics images. Many of the graphics images are built sequentially, one line at a time, and were designed to mimic the manner and speed which the instructor would have used in a chalkboard presentation or a presentation using color pens, transparencies, and an overhead projector. The computer graphics images, however, are in full color, more accurately drawn, and easier to interpret than hand-drawn graphs.

Additional elements are also important in the new teaching method. The text and graphics images displayed on-screen are supplemented by a study guide containing the charts, graphs, and other key lecture materials for the course.² The study guide closely follows the computer-generated slide shows. Gaps are built into the study guide that the instructor "fills in" during the lecture. The study guide is unique because it also functions as the student's notebook. A series of

six comprehensive worksheets containing questions and exercises and based on the study guide and lecture presentation cover every major topic. These worksheets force students to review notes taken during class and the related material in the study guide. Two supplemental computerized slide shows are available to students for individualized instruction. These slide shows provide additional help in completing the exercises dealing with production and cost and are available to students on computers in the undergraduate computer lab. All of these elements and methods are closely interrelated. For example, the study guide is based on the computer-generated lecture images, worksheets are based primarily on study guide and lecture presentations, and the supplemental slide shows are based on the two most difficult worksheets. Examinations and quizzes cover the entire course but place particular emphasis on worksheet material. We evaluated the effectiveness of each of these elements.

The Course

Because of the variety of material covered, the introductory agricultural economics course provided a nearly ideal vehicle for testing the potential and limitations of the approach. The course (called GEN 1C1) consists of three sections. The first section develops an awareness of basic issues confronting U.S. and Kentucky agriculture. (Most of the students enrolled in the course come from rural nonfarm backgrounds or from "part-time" and "hobby" farms, although most are enrolled in the college of agriculture.) Included are issues related to declining farm numbers, changing consumption patterns for agricultural commodities grown in Kentucky, and the importance of various enterprises to farm income in Kentucky. Considerable use is made of charts developed for the course that are similar to those found in the USDA Chartbook, but an emphasis is placed on issues and data important to Kentucky.

The second section introduces basic principles of microeconomic theory using case examples from agriculture. Figures 1 and 2 provide some illustrative examples of the content and level in the microtheory section and illustrate the sequential building of diagrams one line at a time. One of the key elements involves drawing sequences that mimic the manner in which an instructor might draw on a chalkboard or overhead transparency, but with greater clar-

¹ The course was constructed in its entirety using Harvard Graphics, currently used in many agricultural economics departments primarily for research and extension applications drawing technical graphs. A computer disk (for use on an IBM compatible computer with EGA or VGA graphics) containing examples of drawings and lessons developed for the course is available at no charge by contacting either of the authors.

² The study guide contains all of the diagrams and lecture notes displayed on-screen. Some of the graphs that are constructed in several steps and the text material that is presented on-screen one point at a time is included only in final form. The study guide has been duplicated by and sold to students through the campus bookstores. A new version incorporating worksheets and a glossary of terms used in each lesson is available nationwide (Debertin 1990).

A Demand Schedule

Price \$	Quantity Demanded Per Unit of Time
10	0
8	1
7	2
6	3
4	4
3	5
1	6

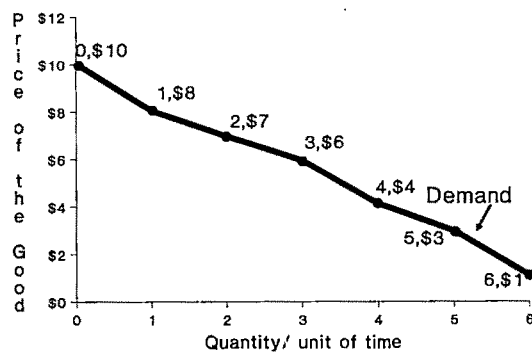
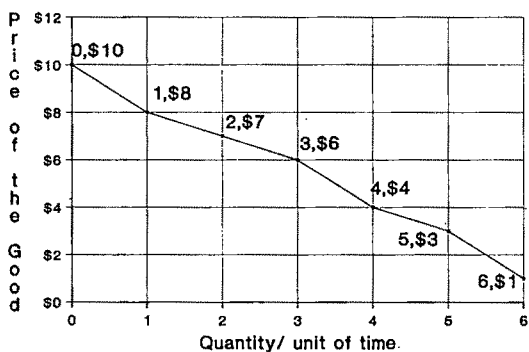
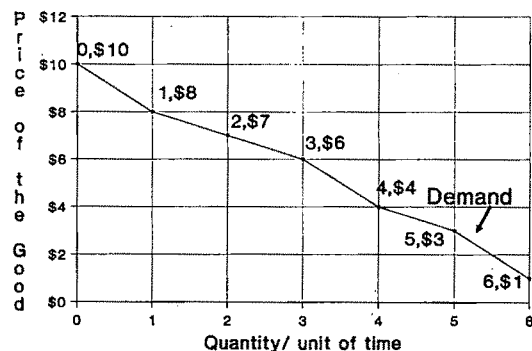
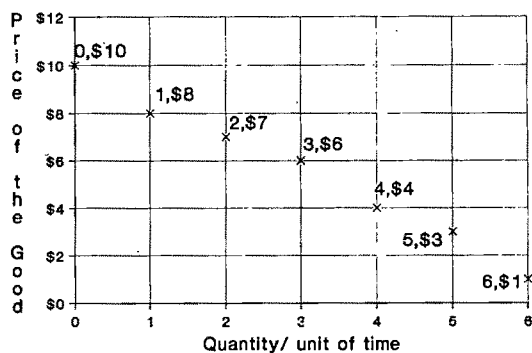
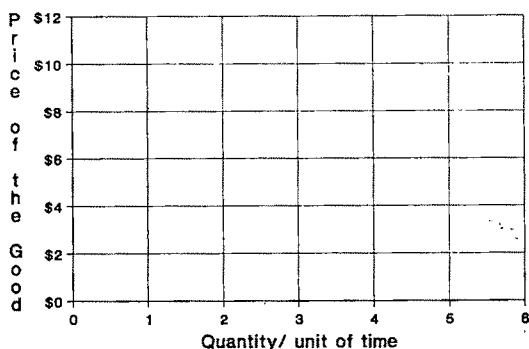


Figure 1. Demand sequence in a microeconomic theory section of the course

ity and accuracy. This careful and accurate drawing is important in a freshman-level class in which students are inexperienced at drawing graphs. For the supply-demand model, the computer is programmed to draw the vertical (price) axis beginning at the top; the horizontal axis from left to right; the demand curve downward from left to right; the supply curve upward from left to right; the equilibrium quantity from the supply-demand intersection downward; the equilibrium price from the supply-demand intersection to the left one line at a time at the same speed and direction as on the chalkboard.

The last section deals with various specializa-

tion areas within agricultural economics, such as marketing, natural resource economics, agricultural policy, and rural development. This section of the course uses a combination of text images, charts to depict data, as well as drawn graphics images similar to those in the micro-theory section.

System Advantages

The system has three primary advantages for the instructor and the students.

(a) *Improved Course Structure and Organi-*

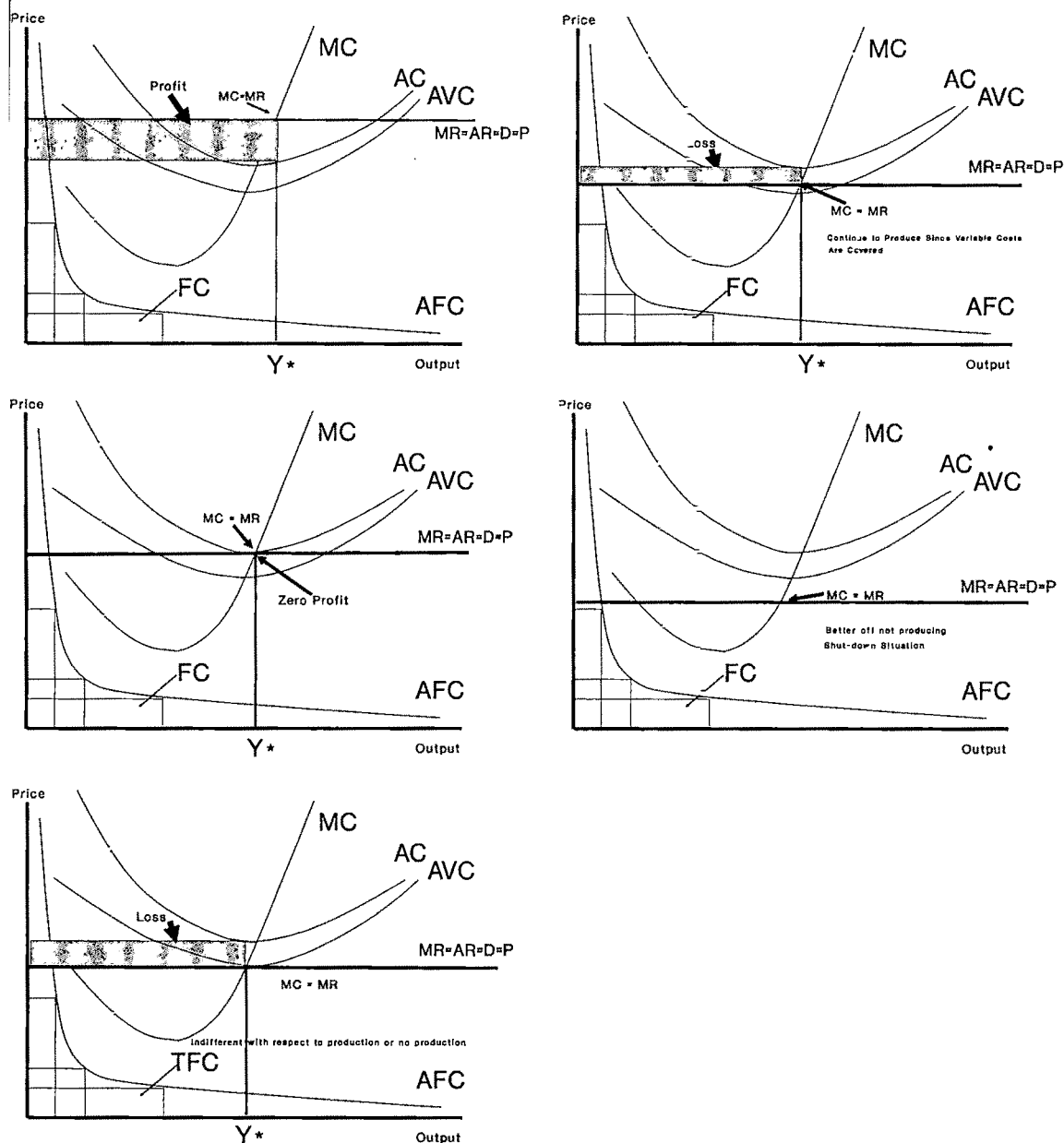


Figure 2. Cost sequence in microeconomic theory section of the course

zational Flow. The flow of the subject matter presentation is logical and sequential. With the combination of study guide and computer-generated slide show, students do not have to determine where a lecture will lead on any given day. If an instructor is absent for a class, another professor can readily substitute without disrupting the flow of the class presentation. If students miss class, they at least know what topics they missed and have the study guide material.

(b) Improved Accuracy of Student Notes. The basic set of notes and graphs from the study guide is correct. Students do not need to worry if they have copied accurately from the chalkboard. This is particularly important advantage for students who miss a lecture.³

³ A concern might be that with access to the detailed study guide, students would tend to miss class more frequently. There is no longer any uncertainty about the topics of any specific lecture. Class attendance is not mandatory. However, we have not observed any

(c) *Improved Use of Class Time.* The course content and depth of coverage has remained the same with the new technologies. However, the utilization of class time has changed somewhat. Now instead of spending valuable class time trying to accurately draw graphs, the instructor can concentrate on explaining the graphs. Instead of copying from the chalkboard, students can concentrate on what the instructor is saying. More emphasis can be placed on interpretation and analysis, with less time devoted to copying the instructor's notes. There is more time to teach processes rather than skills.

Technical Implementation Issues

Computers and Projection Equipment

We have experimented with operation of the course from IBM-compatible computers of various speeds including 8088, 80286, and 80386 machines. With a 80386 machine, special effects, such as line draws mimicking a chalkboard presentation, occur so rapidly that some of the effectiveness is lost. We have found that an 16 Mhz 80286-based machine operates slowly enough that the special video effects realistically mimic how an instructor might draw on the chalkboard and yet fast enough so that the instructor can move through a sequence of screens rapidly, if needed (such as might be desired for a review session). It is not possible to adjust the speed of the draw within the software.

One of the key elements of the technology that makes the approach feasible is the small size (in bytes) of the individual files comprising the individual video screens. Depending on the complexity of the displayed images, these individual files are usually from only 1K–3K. Thus, the entire course needs only 3 megabytes of space on a hard disk. This is because the file comprising a screen image is the data needed to draw the graph, not the actual image. In addition to reducing the space required to storing the screens comprising the course, graphs in the study guide are of high quality, limited only by the resolution of the printer.

All of the lessons are in color. Until recently, we used a monochrome video projector capable of EGA resolution; but in the fall of 1989, we

began using a high-resolution color video projector in a classroom with lighting modified for video display (fig. 3). Video quality from this unit is identical to that from an EGA/VGA computer monitor. We have also experimented with EGA/VGA LCD pads and a conventional overhead projector as a light source, and these units function about the same as the monochrome display projector, with a similar cost. Any of this equipment could be effectively used in the classroom.

Kropp (1973b, p. 760) suggests that "the investment which an academic department wishes to make in instruction is a major factor in determining what teaching methods will be used," and that "computer-based and video courses require heavy initial outlays for producing instructional materials and acquiring equipment." Fortunately, the cost of the equipment we used, though not insubstantial, is declining and is now cheap enough to make investments in the new technologies possible for many, if not most, departments. Cost is one advantage for a monochrome display projector or overhead projector pad (\$2,500 versus \$10,000 for the color projector),⁴ but the line-drawing capabilities and other special video effects are less appealing in monochrome than in color. Most departments undoubtedly already have access to the laser printer needed to produce a high-quality study guide such as the one we developed.

Operational Procedures

The day-to-day equipment operation procedures are new for the instructor, but because the projector and computer are permanently installed in the classroom, less than five minutes are needed to start the equipment before each class. Lessons are copied to the hard disk in the classroom before the semester begins. Computer software is loaded as soon as power is supplied to the computer and projector. The instructor merely selects the "slide show" option from the main program menu and goes to the subdirectory containing the slide show lesson for that day. If necessary, a mouse on the computer functions like a remote control on a slide projector per-

reduction in the proportion of students attending class on any specific day since the system was adopted. In some instances, attendance was up. We estimate that class attendance has averaged about 70% for the semester with the new teaching system.

⁴ Five years ago a color projector capable of EGA/VGA resolution would have cost \$50,000, an amount that would not have been a feasible investment for most agricultural economics departments. Furthermore, a microcomputer capable of generating the high-resolution images would have been considerably more expensive. Most important, easy-to-use and inexpensive software for generating the lessons was not available.

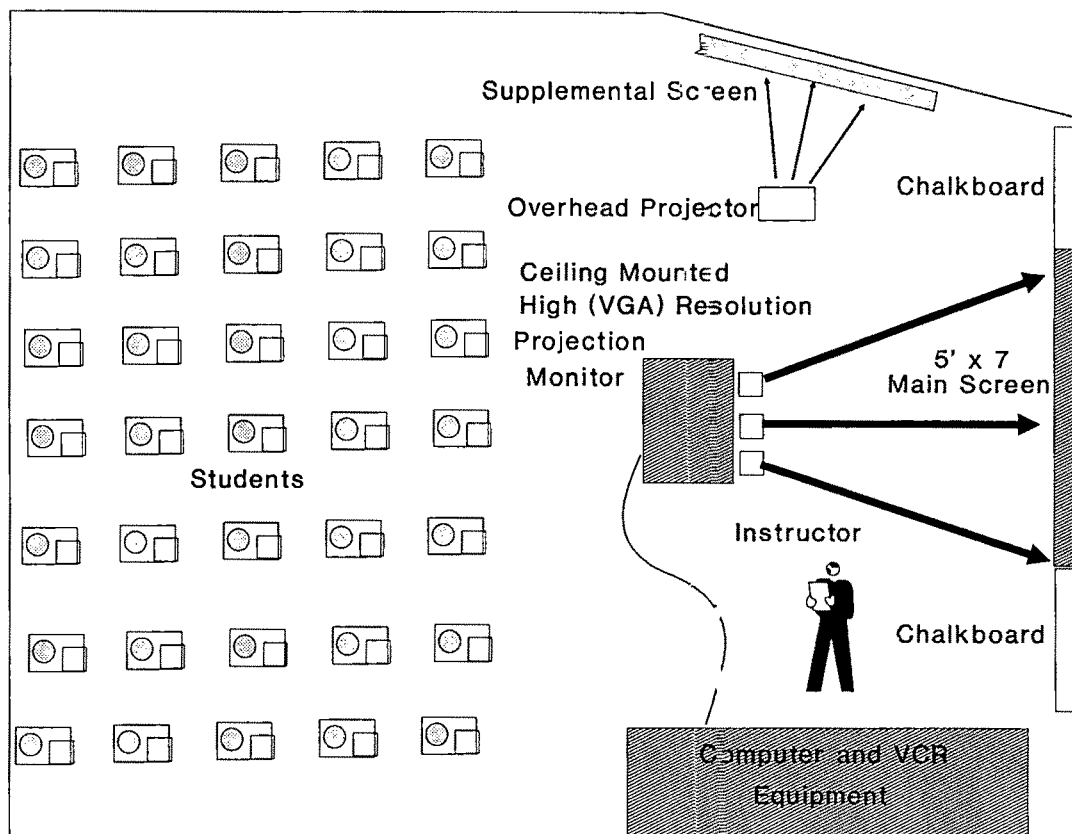


Figure 3. Classroom layout

mitting the instructor to back up during a lecture.

Mixed-media approaches are also possible with the system, and the instructor need not rely entirely on the video display. For example, overhead transparencies are occasionally useful in teaching the course to quickly review a portion of a previous lecture. Any of the 1,200 computer slides comprising the course can be made into a transparency to supplement the video presentation or as a backup system in the event of equipment failure. These transparencies can also be made in color using a pen plotter. An effective classroom presentation also can be made using conventional color slide projectors, as the software is capable of sending text and graphics images to a color slide maker.

Evaluation of Effectiveness

A key concern is the effectiveness of the teaching approach relative to conventional methods of instruction. Do the students prefer the new

approach? Do they learn better? Both instructors have had previous experience teaching the course using hand-drawn transparencies and an overhead projector. During the fall semester, 1989, an evaluation survey was conducted of the GEN 101 students. All but one of the forty-five students completing the class provided responses to the survey. The survey was designed to examine student reaction to the technologies employed as learning aids in the course and to provide students with an opportunity to suggest further course improvements.⁵

Each of the items used in the course (including the new technologies) represents an investment by either the university, the student, or both. This is also true for conventional instructional methods. To illustrate, the purchase of a text-

⁵ In addition to the class evaluation reports reported here, the college also conducts its own evaluation of instructors using a five-point scale. These evaluations focus heavily on the personality of the instructor, how the instructor relates to students, and course organization. Since the new system has been put in place, both instructors have experienced an increase in their average scores on this evaluation of approximately 0.3 to 0.4 of a point.

book is an investment by the student. The computer and display equipment is an investment by the university. Comprehensive worksheets must be graded either by the instructor or a teaching assistant, both of whom are funded by the department. It is important to know which investments yield the greatest effectiveness.

The survey asked students to identify the instructional methods used in the course that most enhanced their learning. The instructional methods evaluated are listed in table 1. Students were asked to use a five-point scale to evaluate the importance of each of these methods in helping them learn the material contained in GEN 101. The components evaluated and the results are presented in table 1. The scale was 5 = very important; 4 = important; 3 = neutral; 2 = unimportant; 1 = very unimportant.

Results of Evaluation

The evaluation occurred during the final week of classes; hence, student reaction to the importance of the final exam in enhancing learning could not be evaluated. Results of this portion of the evaluation are contained in table 1. Based

on the mean values for each response, the following conclusions were reached about the technologies employed.

Students felt strongly about the importance of the worksheets in enhancing learning. This was the first semester we employed worksheets which dealt with all the material contained in the course. Even though answers to questions contained in many of the worksheets could be obtained readily from the study guide, students still felt that the worksheets were the most important learning aid. The study guide also received high marks from students as a learning aid, and students ranked the study guide higher than the computerized slide shows upon which the study guide was based. Based on means from the survey results, we concluded that the animations of the diagrams used in the course made possible by the computer were successful in enhancing learning for the students.

In comparison with the other class materials, the textbook (available to students on reserve in the agricultural library) fared poorly as a learning aid, despite specific assignments related to the lectures and instructor admonitions to read the assignments. Students largely felt that the textbook assignments were redundant. With the

Table 1. Student Response to Learning Effectiveness for Specific Teaching Technologies Used in GEN 101, Fall Semester, 1989

Teaching Technology ^a	Respondent Mean ^b	Standard Error ^c	Percent Ranking as					Did Not Rank
			Very Important				Unimportant	
			5	4	3	2	1	
1. Six worksheets	4.73	0.08	75.0	22.8	2.3	0	0	0
2. Study guide and lecture notes	4.48	0.11	61.4	38.6	13.6	0	0	0
3. Hour exams	3.93	0.13	27.3	43.2	27.3	2.3	0	0
4. The fact that the diagrams "moved" on screen	3.89	0.15	31.8	34.1	27.3	4.5	2.3	0
5. Instructor	3.86	0.12	20.5	52.3	20.4	6.8	0	0
6. Computer-generated lecture slides	3.79	0.16	29.5	29.5	31.8	2.3	4.5	2.3
7. Ten-minute (pop) quizzes	3.50	0.17	18.2	36.4	29.5	9.1	6.8	0
8. The fact that the slides were in color	3.43	0.20	25.0	27.3	27.3	6.8	13.6	0
9. Supplemental slide shows for help on worksheets	3.12	0.20	18.2	18.2	31.8	15.9	13.6	2.3
10. Specific textbook assignments	2.32	0.22	9.1	13.6	20.4	18.2	34.1	4.5
11. Textbook on reserve	2.07	0.22	2.3	6.8	22.7	31.8	36.4	0

^a We would classify technologies on lines 4, 6, 8, and 9 as new technologies; those on lines 3, 5, 7, 10, and 11 as old technologies; and those on lines 1 and 2 as old technologies that have been adapted to take advantage of elements of the "new" technologies. Without the computer-generated slide shows, it would have been difficult to develop the coordinated study guide and set of worksheets.

^b Respondents exclude those who did not rank a technology; $n = 44$.

^c The smaller the standard error, the greater the agreement among respondents with respect to the ranking.

Table 2. Specific Student Rankings of Instructional Components of GEN 101

Instructional Component	Mean Ranking	Standard Error
Worksheets	2.48	0.33
Study guide	3.27	0.40
Computer-generated slides	4.09	0.28
Lecture itself	4.18	0.41
Hour exams	4.52	0.30
Pop quizzes	5.11	0.34
Supplemental slide shows	6.59	0.46
Studying with classmates	6.61	0.49
Old exams	7.14	0.47
Textbook	7.86	0.39

* Students were asked to rank each item from 1 to 10 in terms of its importance in helping them learn the material presented in GEN 101. A 1 indicates a high-ranked item, 10 a low-ranked item. The lower the score, the higher the ranking. The smaller the standard error, the less disagreement among the students with respect to the ranking.

computer slides, study guide, and worksheets, we have been unsuccessful in convincing students of the value in reading from a conventional textbook. However, the results should not be interpreted as a rejection of the use of textbooks in beginning courses. Rather, if the instructor wishes to use a textbook to supplement the written class notes, the instructor needs to determine exactly how the students should use the supplemental textbook material. It is not enough to assume that students will be enthusiastic about seeing another approach to the material contained in the lecture and study guide.

The two supplemental slide shows did not receive high marks as learning aids, even though they were closely linked to, and helpful in,

completing the two most difficult worksheets. The supplemental slide shows also introduced the students to the undergraduate computer lab. However, less than a third of the class indicated that they had gone to the computer lab to view at least one of the slide shows.

Students were then asked to rank ten of the methods with respect to importance in helping them learn the material presented in GEN 101. Results are presented in table 2. Results from the student rankings were largely consistent with the earlier findings. Once again, the worksheets, the study guide, and the computer-generated slides ranked as the top three items; the textbook and old exams ranked lowest.

Finally, students were asked to react to the computer-generated slide shows as a primary mode for lecture presentations. Students could select from four responses or write their own response. These responses are listed in table 3. The strong positive reaction by students to the computer-generated slide shows is evident. Nearly one-half (21 students or 47.7%) indicated that the computer-generated slides were superior to any other means they had seen instructors use to present lecture material. Only two students did not like the slides at all, and one other student responded that the instructor should rely primarily on a chalkboard or an overhead projector instead. Most of the remaining students indicated that, while they preferred the computer-generated slides most of the time, an occasional presentation on the chalkboard or on an overhead would have strengthened the course. These students might have liked one of the "mixed media" approaches.

Table 3. Specific Student Responses to Computer-Generated Slide Shows

- The computer-generated slides were superior to any other means I have seen instructors use to present lecture material.*
21 selected this response, representing 47.7% of the class.
- The computer-generated slides were OK, but sometimes I wished that the instructor would use a chalkboard or a visual on an overhead projector to make a point.*
13 selected this response, representing 29.5% of the class.
- I wish that the instructor would rely primarily on the chalkboard or an overhead projector for his presentations. The computer is OK to supplement the lectures on some material, but should not be used everyday.*
1 person selected this response, representing 2.3% of the class
- The computer-generated slides were a bomb. Why can't we have ordinary lectures in GEN 101?*
2 students selected this response, representing 4.5% of the class
- None of the above statements adequately represents my response.*
6 students, representing 13.6% of the class provided other responses. Most of these were qualified endorsements similar to statement number 2.

Suggestions for Course Improvement

In order to identify possible improvements in the course, students were asked, again on the basis of a five-point scale, whether they agreed or disagreed with a series of statements each suggesting a change in the course. Results are presented in table 4. Particular emphasis was placed on eliciting responses indicating dissatisfaction with the computer-generated slide shows in favor of returning to a conventional presentation mode. Most of the students were not in favor of going back to the chalkboard and overhead projector. Students again indicated that the study guide based on the computer-generated slides was adequate without a textbook. The suggestion with the strongest support was for creating a file of old exams. Other suggestions receiving considerable support among students were for spending more time on class discussion and going over exams, quizzes, and worksheets.

Impact on Student Performance on Exams

An effort was made to determine if students performed better on exams with the new system in place. Exam scores for the GEN 101 class were compared for the fall semester 1987, before the

system had been implemented and fall semester 1989, when the complete system was used. A number of qualifications are needed when making such student performance comparisons. There is no assurance that the students enrolled in the class in the fall 1989 were identical in ability to those enrolled in the fall of 1987. Another difference was that the enrollment in the fall of 1987 was somewhat higher than in the fall of 1989 (68 versus 45 students). However, the instructor was the same in the fall of 1987 as in the fall of 1989, the course content in terms of topics covered and level of difficulty was similar. Examinations were also similar in structure and level of difficulty in both years and consisted of a combination of true and false and multiple choice questions, short answer essays, problems, and questions that required students to draw graphs.

Table 5 reports the mean exam scores for two hour exams and the final exam under the old and new system. Mean exam scores were significantly higher for the first hour exam (83.29 in 1989 versus 74.31 in 1987) and for the final exam (82.44 in 1989 versus 74.66 in 1987). No significant difference was found for the second hour exam (72.60 in 1989 versus 74.22 in 1987). These data provide at least a small amount of evidence to support the hypothesis that the system had a positive impact on learning.

Table 4. Specific Student Suggestions for Improving GEN 101, Fall Semester, 1989

Suggestion	Mean	Standard Error	Percent Responding				Strongly Disagree 1	No Response
			Strongly Agree 5	4	Neutral 3	2		
Create a file of old exams	4.09	0.16	43.2	31.8	20.4	0	4.5	0
Study guide adequate without book	4.05	0.18	47.7	20.5	20.5	4.5	2.3	2.3
Spend more time going over worksheets	3.43	0.15	11.4	43.2	27.3	13.6	4.5	0
More class discussion is needed	3.30	0.14	11.4	25.0	47.7	13.6	2.3	0
Spend more time going over exams and quizzes	3.23	0.16	15.9	15.9	47.7	15.9	4.5	0
More worksheets should be used	2.98	0.18	15.9	9.1	38.6	29.5	6.8	0
Use the chalkboard more	2.77	0.17	9.1	11.4	38.6	29.5	11.3	0
Spend more time going over supplemental slide shows	2.70	0.17	4.5	20.5	31.8	22.7	18.2	2.3
Pop quizzes should not be used	2.67	0.21	13.6	9.1	27.3	22.7	22.7	4.5
Use an overhead projector more	2.61	0.19	13.6	6.8	22.7	40.9	15.9	0
Use computer-generated slides less	2.05	0.15	4.5	2.3	18.2	43.2	31.8	0
Should require textbook purchase	1.93	0.19	6.8	6.8	11.4	20.5	52.3	2.3
Worksheets need not be graded	1.61	0.11	0	13.6	34.1	52.3	0	0
Quiz and test questions should come from the textbook on reserve	1.52	0.12	0	2.3	11.4	22.7	63.6	0
Worksheets should not be used	1.52	0.14	0	6.8	6.8	18.2	68.2	0

Table 5. Differences in Mean Exam Scores Fall 1987 and Fall 1989

<u>First hour exam:</u>				
	N	Mean	Standard Deviation	
1989	45	83.29	12.42	
1987	68	74.31	17.63	
Variances assumed	<i>t</i>	<i>DF</i>	Probability > <i>t</i>	
Equal (pooled)	2.96	111	0.0037	
Unequal (separate)	3.18	110.5	0.0018	
<i>F</i> -test for equal variances: <i>F</i> = 2.016 with <i>DF</i> 67, 44. Probability > <i>F</i> = 0.0146				
<u>Second hour exam:</u>				
	N	Mean	Standard Deviation	
1989	45	72.60	12.38	
1987	68	74.22	14.63	
Variances assumed	<i>t</i>	<i>DF</i>	Probability > <i>t</i>	
Equal (pooled)	-0.61	111	0.542	
Unequal (separate)	-0.63	104.4	0.529	
<i>F</i> -test for equal variances: <i>F</i> = 1.396 with <i>DF</i> 67, 44. Probability > <i>F</i> = 0.240				
<u>Final exam:</u>				
	N	Mean	Standard Deviation	
1989	45	82.44	14.14	
1987	68	74.66	12.26	
Variances assumed	<i>t</i>	<i>DF</i>	Probability > <i>t</i>	
Equal (pooled)	3.11	111	0.002	
Unequal (separate)	3.02	84.9	0.003	
<i>F</i> -test for equal variances: <i>F</i> = 1.33 with <i>DF</i> 67, 44. Probability > <i>F</i> = 0.288				

Concluding Comments

It is both possible and practical to develop a course based on computer graphics as the primary medium of instruction employing currently available computer, electronics-imaging, and software technologies. The survey results revealed that the combination of computer-generated slides along with a coordinated study guide and set of worksheets provide learning opportunities beyond a conventional chalkboard presentation supplemented with textbook assignments. Survey results provide strong evidence that the students prefer the new teaching approach. In addition, some evidence suggests that students learned more when the innovations were employed.

The development of the video screens for the lecture presentation and study guide was a part-time effort by one of the instructors over a six-month period. Although the software used is simple to use, the speed at which the instructor can create the video images improves significantly with practice. The last lessons in the course were developed far more quickly, and with greater ease and accuracy than were the first lessons. Once the instructor has become skilled with

the capabilities of the software, a typical lecture can be developed in perhaps two hours. Only lessons involving the use of graphs based on secondary data take longer. This is not inconsistent with the amount of time an instructor would ordinarily use to prepare a lecture for a freshman-level course employing conventional techniques.

[Received September 1989; final revision received May 1990.]

References

- Bay, Ted F., and Richard A. Schoney. "Data Analysis with Computer Graphics: Production Functions." *Amer. J. Agr. Econ.* 64(1982):289-97.
- Beck, Robert L., A. Frank Bordeaux, Jr., Joe T. Davis, Russe H. Brannon, and Loys L. Mather. "Undergraduate Programs in Agricultural Economics: Some Observations." *Amer. J. Agr. Econ.* 59(1977):766-68.
- Bentley, Ernest. "Computer Use in the Agricultural Economics Classroom: Discussion." *Amer. J. Agr. Econ.* 64(1982):936-87.
- Brown, Thomas G. "Changing Delivery Systems for Agricultural Extension: The Extension Teacher—Changing Roles and Competencies." *Amer. J. Agr. Econ.* 63(1981):859-62.
- Coffey, Joseph D. "Undergraduate Agricultural Economics Curricula: Discussion." *Amer. J. Agr. Econ.* 69(1987):1043.
- Debertin, David L. *Agricultural Production Economics*. New York: Macmillan Co., 1986.
- . "Developing Realistic Production Functions for Use in Undergraduate Classes." *S. J. Agr. Econ.* 17(1985):207-14.
- . *Economics of Food and Agriculture: A Study Guide*. Dubuque IA: Kendall Hunt, 1990.
- . "Economics of Food and Agriculture," class notes for GEN 101, Dep. Agr. Econ., University of Kentucky, 1988 and (2nd ed.) 1989.
- Debertin, David L., Angelos Pagoulatos, and Garnett L. Bradford. "Computer Graphics: An Educational Tool in Production Economics." *Amer. J. Agr. Econ.* 59(1977):573-76.
- Diesslin, H. G. "The Computer—Extension's Delivery System of the Future." *Amer. J. Agr. Econ.* 63(1981):863-67.
- Dobson, William D. "Undergraduate Agricultural Economics Curricula: Discussion." *Amer. J. Agr. Econ.* 69(1987):10-5.
- Erven, Bernard L. "Reforming Curricula: Challenge and Change for Agricultural Economists." *Amer. J. Agr. Econ.* 69(1987):1037.
- Fuller, Earl L. "Microcomputers: Useful in All of Agricultural Economics and Extension." *Amer. J. Agr. Econ.* 64(1982):978-85.
- Harris, Thomas F. "Microcomputer Graphics for a Commodity Futures Course." *Proceedings of the AAEA*

- Teaching Workshop, Reno, NV*, 1986, pp. 83–88.
- Holt, John. "Changing Delivery Systems for Agricultural Extension: Discussion." *Amer. J. Agr. Econ.* 63(1981): 868–69.
- Hudson, Michael A., Raymond M. Leuthold, Clark A. Roberts, Brian D. Adam, and Steven T. Sonka. "Microcomputer-Based Networks for Teaching Agricultural Marketing." *Proceedings of the AAEA Teaching Workshop, Reno, NV*, 1986, pp. 176–81.
- Hughes, Harlan. "Changing Delivery Systems for Agricultural Extension: Discussion." *Amer. J. Agr. Econ.* 63(1981):870.
- Kropp, Russell P. "Curriculum Development: Principles and Methods." *Amer. J. Agr. Econ.* 55(1973a):735–39.
- . "Teaching Method." *Amer. J. Agr. Econ.* 55(1973b): 757–61.
- Litzenberg, Kerry K. "Computer Use in the Agricultural Economics Classroom." *Amer. J. Agr. Econ.* 64(1982): 970–77.
- Manderscheid, Lester V. "Guidelines for Curriculum Changes in Agricultural Economics." *Amer. J. Agr. Econ.* 55(1973):740–47.
- Mather, Loys L., Joe T. Davis, Russel H. Brannon, A. Frank Bordeaux, Jr., and Robert L. Beck. "Developing a Competency-Based Curriculum in Agricultural Economics." *Amer. J. Agr. Econ.* 59(1977):760–65.
- Osburn, D. D., K. C. Schneeberger, M. R. Wilsdorf, and E. S. Reber. "Microcomputer Aided Instruction." *N.A.C.T.A. Journal* 25(1981):24–26.
- Wetzstein, Michael. "An Organonic and Modern Problems Approach for Teaching Agricultural Economics Principles." *Amer. J. Agr. Econ.* 70(1988):63–67.

Supply Response Under Proportional Profits Taxation

John Quiggin

A striking result in the theory of the competitive firm under certainty is the proposition that a proportional profits tax (with full offsets for losses) will have no impact on optimal output. This result does not apply under uncertainty. It is shown that, under constant or increasing returns to scale, a proportional profits tax will yield an unambiguous expansion in output. The same result is shown to hold for the more general rank-dependent expected utility (RDEU) model.

Key words: firm scale, risk, uncertainty.

A striking result in the theory of the competitive firm under certainty is the proposition that a proportional profits tax (with full offsets for losses) will have no impact on optimal output. The nondistortionary properties of such a tax have been the basis of a large literature incorporating attempts to devise a real tax which approximates the theoretical ideal as closely as possible (e.g., Brown).

This result does not apply under uncertainty. In his seminal expected utility analysis of the firm under uncertainty, Sandmo claimed that a proportional profits tax would increase or decrease output according to whether the coefficient of relative risk aversion is increasing or decreasing. Katz pointed out that this result was erroneous and proposed several complex conditions involving both absolute and relative risk aversion. In their recent survey of the theory of the firm under uncertainty, Robison and Barry use a mean-variance approximation to the expected utility (EU) model to examine this question and conclude that the impact of a proportional profits tax is ambiguous except in the restrictive special case of constant absolute risk aversion.

The competitive optimal output under certainty is only defined when returns to scale are, at least eventually, decreasing. This situation, however, does not apply under uncertainty.

Quiggin (1982a) shows that, under weak assumptions, a competitive optimal output will exist under constant or increasing returns to scale.¹ This analysis is particularly important in relation to agriculture because many studies have found evidence of increasing returns to scale. The most usual result is the existence of an L-shaped cost curve with an initial region of increasing returns followed eventually by constant returns. (A summary of the literature, along with a critical response, is given by Castle.) Risk aversion is a promising explanation for the observed existence of competitive firms of finite size.

The object of this note is to show that, under constant or increasing returns to scale, a proportional profits tax will yield an unambiguous expansion in output. The same result is shown to hold for the more general rank-dependent expected utility (RDEU) model (Quiggin 1982b, Yaari). The result does not generalize to other general models such as the smooth preferences model of Machina.

The seemingly perverse supply response derived here may seem less surprising when it is considered in the light of the debate over the treatment of risk in public projects following the work of Arrow and Lind. In this literature, the analysis is cast in terms of discount rates. However, there is a direct relationship to the problem

John Quiggin is a visiting associate professor in the Department of Agricultural and Resource Economics, University of Maryland.

The author would like to thank Darrell Hueth and especially Bob Chambers, as well as two anonymous referees, for helpful comments and criticism. They are, of course, not responsible for any remaining errors.

¹ The term returns to scale is sometimes used in the specific sense of the cost structure in relation to a proportionate expansion in all inputs, and distinguished from the case, referred to as returns to size where the least-cost input mix varies as output increases. In this analysis, as in most of the literature on the firm under uncertainty, the term returns to scale will be used in the general sense including returns to size.

of the firm under uncertainty because the more the discount rate is increased to take account of risk, the lower is the level of investment and output. One of the themes in that literature is that a proportionate corporate profits tax makes the government an effective partner in private projects, bearing a share of the risk proportionate to the tax rate (Mayshar). This relationship implies that the discount rate for such projects should be lowered; that is, the number of projects undertaken should be larger than it would be for the same market rate of interest in the absence of a corporate tax rate. Thus, at least in a partial equilibrium sense, this analysis leads us to expect a positive supply response to a proportional profits tax. A similar analysis applies to risk-sharing problems such as those of sharecropping and franchise contracts.

The positive supply response result may also be motivated by considering the case of constant returns to scale. Under constant returns to scale, for any realization of the random price variable p , profit is linear in output α . It is well known (Sandmo, Quiggin 1982a) that, whereas output under certainty is indeterminate in this case, a risk-averse firm under uncertainty will have a finite optimal output. Denote this output α^* and let the associated distribution of profits be $F(\pi|_{\alpha^*})$. Now suppose a tax is imposed at a rate of 50%. By the linearity property of constant returns to scale technology, the distribution $F(\pi/2|_{2\alpha^*})$ obtained at an output of $2\alpha^*$ is identical to the original optimum $F(\pi|_{\alpha^*})$ and is the new optimal distribution.

The Economic Control Problem

The positive supply response result will be proved in the context of a general economic control problem which includes as special cases the problem of the firm under price or yield uncertainty. The firm is required to solve the maximization

$$(1) \quad \text{Max}_{\alpha} E[U(\pi(\theta, \alpha))],$$

where θ is an economically relevant random variable, such as price or yield, and α is a control variable, interpreted here as (planned) output. The profit function π determines the level of profits for given choices of θ and α . In EU theory, the function U is a von Neumann-Morgenstern utility function, here assumed to be concave. It will be assumed that

$$(A.1) \quad \partial \pi / \partial \theta > 0, \text{ and}$$

$$(A.2) \quad \partial^2 \pi / \partial \theta \partial \alpha \geq 0.$$

These conditions will be fulfilled in all the standard versions of the control problem. The first ensures that increasing values of the random variable (e.g., output price) increase profits. This condition is a natural one in most versions of the problem. The second condition ensures that increasing values of the random variable increase the marginal return to the control variable (e.g., output). For the firm under output price uncertainty, we have

$$(2) \quad \pi(\theta, \alpha) = \theta \alpha - C(\alpha),$$

and (A.2) is clearly satisfied. Many more general forms of the profit function will also satisfy (A.2).

The first-order condition for an interior optimum for the problem (1) is

$$(3) \quad E[\partial U / \partial \pi \partial \pi / \partial \alpha] = 0,$$

and the second-order condition is

$$(4) \quad D = E[\partial^2 U / \partial \pi^2 (\partial \pi / \partial \alpha)^2 + \partial U / \partial \pi \partial^2 \pi / \partial \alpha^2] < 0.$$

Given the assumed concavity of U , the second-order condition will always be satisfied if π is concave or linear in α .

It remains to be shown that an interior solution will exist. The decision maker must be sufficiently risk averse to permit the existence of a finite optimal output, while production must be sufficiently attractive that the optimal output is nonzero. Quiggin (1982a) derives necessary and sufficient condition for the existence of a finite optimum. It is assumed that there is, in the limit, a positive probability that increases in the control variable will reduce profit. That is,

$$\lim_{\alpha \rightarrow \infty} \text{Pr}\{\partial \pi / \partial \alpha \leq 0\} > 0.$$

Given this "no easy money" assumption, a fairly weak sufficient condition on the utility condition is $\lim_{\pi \rightarrow \infty} U'(\pi) = 0$. This condition is satisfied by all the constant relative risk-aversion functions and, hence, by any function in which the coefficient of relative risk-aversion is bounded away from zero.

The second requirement is that the optimal output should be positive. A sufficient condition may be derived from the observation that, for small risks, the decision maker will be effectively risk neutral. Thus, if $E[\partial \pi / \partial \alpha]$ is positive in a neighborhood of $\alpha = 0$, a positive level of output will be chosen.

Now let a proportional profit tax, with full rebates for losses, be imposed at a rate τ . Then the problem becomes

$$(5) \quad \text{Max}_{\alpha} E[U(\pi_0(\theta, \alpha))],$$

where $\pi_0 = (1 - \tau)\pi$, with first- and second-order conditions

$$(6) \quad E[\partial U / \partial \pi_0 \partial \pi_0 / \partial \alpha] = 0, \text{ and}$$

$$(7) \quad D = E[\partial^2 U / \partial \pi_0^2 (\partial \pi_0 / \partial \alpha)^2 + \partial U / \partial \pi_0 \partial^2 \pi_0 / \partial \alpha^2] < 0.$$

Implicit differentiation of (6) with respect to τ yields

$$(8) \quad \partial \alpha^* / \partial \tau = (1/D) E[\pi \partial^2 U / \partial \pi_0^2 \partial \pi_0 / \partial \alpha],$$

where, as before, α^* denotes the optimum level of α .

By (A.1), for any α , there is a unique value $\bar{\theta}$ such that $\theta \geq \bar{\theta} \Leftrightarrow \partial \pi / \partial \alpha \geq 0$. Let $\bar{\pi}$ be the value of π given $\theta = \bar{\theta}$; then,

$$(9) \quad \partial \alpha^* / \partial \tau = (1/D) \{ E[(\pi - \bar{\pi}) \partial \pi_0 / \partial \alpha \partial^2 U / \partial \pi_0^2] + \bar{\pi} E[\partial \pi / \partial \alpha \partial^2 U / \partial \pi_0^2] \}.$$

The first expectation on the right-hand side (RHS) is everywhere negative by the definition of $\bar{\pi}$, since $(\pi - \bar{\pi})$ has the same sign as $\partial \pi_0 / \partial \alpha$ for every value of θ . The second expectation may be shown to be positive as follows. Observe that $\partial^2 U / \partial \pi_0^2 = -A \partial U / \partial \pi_0$, where A is the coefficient of absolute risk aversion. Thus,

$$(10) \quad E[\partial \pi / \partial \alpha \partial^2 U / \partial \pi_0^2] = -E[A \partial \pi_0 / \partial \alpha \partial U / \partial \pi].$$

From the first-order condition (6), the second expectation in the RHS of (9) will be zero whenever A is constant. This yields the result, derived by Robison and Barry, that $\partial \alpha^* / \partial \tau \geq 0$ under constant absolute risk aversion.

For the more plausible case of decreasing absolute risk aversion, we may proceed as follows. Let \bar{A} be the value of A when $\theta = \bar{\theta}$. Because A is monotone decreasing, while $\partial \pi_0 / \partial \alpha \partial U / \partial \pi$ is positive if and only if $\theta \geq \bar{\theta}$,

$$(11) \quad E[A \partial \pi / \partial \alpha \partial U / \partial \pi] \leq \bar{A} E[\partial \pi / \partial \alpha \partial U / \partial \pi] = 0,$$

by (6).

Thus, if $\bar{\pi}$ is negative, the entire term in curly brackets in equation (9) is negative. Then, by (8), $\partial \alpha^* / \partial \tau \geq 0$.

When the control problem is that of the competitive firm under output price uncertainty, θ

represents price and the definition of $\bar{\theta}$ implies $\bar{\theta}$ is equal to marginal cost. Profit is equal to output multiplied by price less average cost. Thus, the sign of $\bar{\pi}$ depends on whether $MC < AC$. Under constant returns to scale, $MC = AC$ and $\bar{\pi} = 0$. Under increasing returns to scale, $MC < AC$ and $\bar{\pi} < 0$.

The discussion above yields

PROPOSITION 1. *For the competitive firm under output price uncertainty, given decreasing absolute risk aversion, a sufficient condition for a positive supply response to an increase in the rate of proportional profit taxation is that returns to scale should be nondecreasing.*

For a U-shaped cost curve displaying eventually decreasing returns, the same argument establishes that there will be a positive supply response whenever the initial optimum is in the region of the curve characterized by declining or flat average costs; that is, whenever output is at or below the point at which average costs are minimized. A production point below the cost-optimizing level can never be optimal for the competitive firm in the absence of uncertainty, but it is perfectly feasible when uncertainty is present.

The extension to RDEU preferences is straightforward. Maximization of an RDEU functional is equivalent to maximization of expected utility with respect to a transformed probability distribution. For a problem such as that of a proportional profits tax, therefore, the EU result carries over directly. This technique does not work for the general smooth preferences modelled by Machina. The change in the distribution of income associated with a proportional profits tax will lead to a change in the local utility function with respect to which the first-order condition (6) applies, and this ensures that comparative static results will be ambiguous.

Concluding Comments

The existence of a tax which calls forward a positive supply response might seem like the discovery of a fiscal "philosopher's stone."² However, the analysis presented here does not apply to ordinary forms of capital taxation, such

² The discovery of the philosopher's stone, which would transmute base metals into gold, was a key research objective of the medieval alchemists.

as corporate income taxes. The tax base is pure economic profit, that is, above-normal returns to capital and the tax must include full offsets for firms with below-normal returns to capital. Hence, if a proportional profits tax of the type analyzed here were applied across the board, the net revenue raised would be unlikely to be large and might not even be positive.

The appropriate role for taxes of this kind arises when the tax base is some form of economic rent. As noted in the introduction, proportional profits taxes have been proposed as a means of taxing the rent associated with mineral deposits. A similar approach might be used as a means of financing agricultural developments, such as irrigation schemes, particularly where the returns to farmers are very risky. Rather than selling water rights at market prices, governments could allocate them at low cost and impose a proportional levy on (pure economic) profit. In both mining and agricultural contexts, however, numerous practical difficulties must be overcome before a tax of this kind could be implemented.

*[Received December 1989; final revision
received April 1990.]*

References

- Arrow, K., and R. Lind. "Uncertainty and the Evaluation of Public Investment Decisions." *Amer. Econ. Rev.* 60(1970):364-78.
- Brown, E. C. "Business-Income Taxation and Investment Incentives." *Essays in Honor of Alvin Hansen*. New York: W. W. Norton Co., 1948.
- Castle, E. "Is Farming a Constant Cost Industry?" *Amer. J. Agr. Econ.* 71(1989):574-82.
- Katz, E. "Relative Risk Aversion in Comparative Statics." *Amer. Econ. Rev.* 73(1983):452-53.
- Lin, E., and E. Katz. "Measures of Risk Aversion and Comparative Statics of Industry Equilibrium." *Amer. Econ. Rev.* 76(1986):524-29.
- Machina, M. "'Expected Utility' Analysis without the Independence Axiom." *Econometrica* 50(1982):277-323.
- Mayshar, J. "Should the Government Subsidize Risky Private Projects?" *Amer. Econ. Rev.* 67(1977):20-28.
- Quiggin, J. "A Note on the Existence of a Competitive Optimum." *Econ. Recd.* 55(1982a):174-76.
- . "A Theory of Anticipated Utility." *J. Econ. Behavior and Organization* 3(1982b):323-43.
- Robison, L., and P. Barry. *The Competitive Firm's Response to Risk*. New York: Macmillan Co., 1987.
- Sandmo, A. "The Competitive Firm Under Output Price Uncertainty." *Amer. Econ. Rev.* 61(1971):65-73.
- Yaari, M. "The Dual Theory of Choice Under Risk." *Econometrica* 55(1987):95-115.

An Econometric Model of the U.S. Beekeeping Industry

Lois Schertz Willett and Ben C. French

This paper presents a dynamic econometric model of the U.S. beekeeping industry for policy analysis and economic projections. Data from 1952 to 1984 were used to estimate the model by three-stage least squares. The model indicates that, when the federal price support for honey exceeded the market price, the federal program had a significant impact on several sectors of the industry. A comparison of beekeeper revenue, consumer expenditures, and federal government expenditures suggests the honey support program was an ineffective means of supporting honey prices from 1982 through 1984.

Key words: beekeeping, dynamic, econometric, honey support program, price analysis.

The U.S. beekeeping industry contributes to the nation's food supply by producing honey and by providing pollination services to over 3.5 million acres of fruits, vegetables, oilseeds, and legume seed crops dependent on insect pollination. Insect pollination increases yields on another 63 million acres (Hoff and Phillips). Bees produced 150 million pounds of honey, and pollination services benefited crops with a total value of \$9.3 billion in 1985 (Robinson, Nowogrodzki, and Morse). Beekeeper returns have been affected by government price support policies, imports, and pesticide applications, and potentially may be affected by the spread of Africanized bees. Previous economic studies of the industry by Meade, Cheung, Johnson, Gould, Siebert, Olmstead and Wooten have focused on more limited issues pertaining to externalities and the effects of bees on productivity. McDowell evaluated the impacts of a possible spread of Africanized bees, and the General Accounting Office evaluated the federal price support program for honey; but neither study formulated a model involving estimated supply and demand relationships. Policy analysis and economic projections may be improved by a more complete quantitative modeling framework.

Lois Schertz Willett is an assistant professor of agricultural economics, Cornell University. Ben C. French is a professor of agricultural economics, University of California, Davis.

This research was supported in part by the Giannini Foundation of Economics.

The authors thank two anonymous reviewers for helpful suggestions.

This paper presents a dynamic econometric model of the beekeeping industry. The model is used primarily to gain insight into the economic effects of changing the price support program for honey. It has further potential as a tool in analyzing economic issues involving (a) the possible effects of the Africanized honey bee; (b) the potential effects of bee diseases and parasites such as chalkbrood, tracheal mites, and the varroa mite; (c) the probable effects of pesticides on the industry; and (d) the import market. The theoretical foundations for modeling the production and marketing in the beekeeping industry are presented first. In the next sections the empirical model is defined, and empirical results are presented. Finally, the impacts of a change in the federal support program are discussed.

Conceptual Model

The primary products of the bee industry are honey, beeswax, and pollination services. The primary production inputs are packages of bees and queen bees, transportation services, extraction and handling equipment, and labor. Some beekeepers generate replacement or expansion colonies from their own brood stock. Other beekeepers, especially in colder climates, purchase replacement packages and queens from beekeepers in warmer climates who specialize in producing these products for sale. The theoretical foundations for the equation system re-

quired to model the production and marketing of bee products are outlined briefly in this section.

Supply Response

A generalized production function for beekeepers is

$$(1) \quad F(H, W, S, C, K, L) = 0,$$

where H , W , S are rates of output of honey, wax (a by-product), and pollination services, C is a bundle of inputs associated with (and measured by) the number of colonies, and K and L are other capital and labor inputs. For firms that replace or expand part or all of their colonies with purchased bees and queens, C may be decomposed into a component proportional to the purchase of package bees (B) and queens (Q) and a component associated with internally generated bee replacements. For firms that produce bees and queens for sale, B and Q are added as products.

A generalized net revenue function for beekeeping may be expressed as

$$(2) \quad R = H \cdot P_H + W \cdot P_W + S \cdot P_S + B \cdot P_B + Q \cdot P_Q - C \cdot P_C - K \cdot P_K - L \cdot P_L,$$

where P_H , P_W , P_S , P_B , and P_Q are producer prices for honey, beeswax, pollination services, packages of bees, and queens, respectively; P_C reflects the direct cost of replacing or adding a colony; and P_K and P_L are vectors of prices of other capital and labor inputs. For beekeepers who purchase packages and queens, P_C is a function of P_B and P_Q and transportation charges. The values of B and Q are zero for firms that do not produce packages or queens, and S is zero for firms that do not sell pollination services.

The determination of optimal input and output rates is a two-stage process. Assuming competitive conditions, the first stage involves maximizing expected net revenue over a future time horizon with respect to all output and input choices, subject to the production function. The first stage generates a set of functions which relate planned inputs and planned outputs to the expected prices of all inputs and outputs. It also provides an input demand function for bee colonies which relates the investment in colonies to expected future prices of all products and inputs.

Changes in the price support program for honey could affect beekeepers' price expectations and,

therefore, the profit-maximizing adjustments. Honey price supports have been in place for many years, but generally they were below producer average market prices prior to 1981. Thus, the price support program would have little effect on producer decisions over that period. From 1981 to the end of the data used for estimation (1984), support prices exceeded free market prices. This change could have affected producer investment response and other equations of the system as well. To reflect this effect, we introduced a variable, X , which is zero prior to 1981 and thereafter represents the difference between the support price and the market price. This specification is preferable to using a zero-one dummy variable because the response likely varies with the difference between support and market prices.

With these considerations, the aggregate colony response function is expressed as

$$(3) \quad C = C(P_H^e, P_W^e, P_S^e, P_B^e, P_C^e, P_K^e, P_L^e, X),$$

where the superscript e indicates an expected value.¹ The producer expectation model is discussed with the empirical specifications.

The current-year outputs of bee products (in contrast to the planned outputs determined in the first stage) flow recursively from investments in bee colonies. Second-stage solutions for the bee product mix are obtained by maximizing current net revenue with respect to all outputs and inputs, subject to the production function with colonies (C) given. The result is a set of conditional short-run supply functions plus input demand functions for packages and queens, which express quantities supplied (or demanded) as functions of colony numbers and current product and input prices and the influence of the price support program. That is,

$$(4) \quad H = H(C, P_H, P_W, P_S, P_B, P_Q, P_K, P_L, X),$$

$$(5) \quad W = aH,$$

$$(6) \quad S = S(C, P_H, P_W, P_S, P_B, P_Q, P_K, P_L, X),$$

$$(7) \quad B^S = B^S(C, P_H, P_W, P_S, P_B, P_Q, P_K, P_L, X),$$

$$(8) \quad Q^S = Q^S(C, P_H, P_W, P_S, P_B, P_Q, P_K, P_L, X),$$

$$(9) \quad B^D = B^D(C, P_H, P_W, P_S, P_B, P_Q, P_K, P_L, X),$$

¹ Investments in bee colonies could be influenced by changes in alternative uses of beekeeper resources. However, we could not identify a clear measure of returns to such alternatives.

and

$$(10) \quad Q^D = Q^D(C, P_H, P_W, P_S, P_B, P_Q, P_K, P_L, X).$$

Equations (4) to (8) are conditional supply curves, and equations (9) and (10) are the input demand functions for bee packages and queens.

Allocation of Honey between Processing and the CCC

Beekeepers forfeit their honey to the Commodity Credit Corporation (CCC) when the support price exceeds the market price. However, the allocation does not jump from 0% to 100%, but it increases with increases in the ratio of the support price to the market price. That is,

$$(11) \quad H_C = A(H, SP, P_H) \text{ if } SP \geq bP_H, \\ H_C = 0 \text{ if } SP < bP_H,$$

where SP is the support price and H_C is quantity sold to the CCC. The coefficient b is a value slightly less than 1.0 since some honey may be allocated to the CCC even when the market price is slightly below the support price because of market imperfections, differences in availability and accessibility of CCC storage facilities across the country, and differences in honey quality characteristics. The quantity sold to processors (H_P) is the difference between the quantity produced and the quantity forfeited to the CCC ($H_P = H - H_C$).

Processor Demand for U.S.-Produced Honey

Honey processors obtain raw honey from both U.S. producers and from other countries. Imported honey is not a perfect substitute for U.S. produced honey because of market contracts, concern over dependence on imports, and variations in quality and type of honey. Under the competitive conditions assumed here, the profit-maximizing behavior of processors generates a set of input demand functions which relate the quantity of each input to the expected price of the processed product and the prices of all inputs. The variable X allows for a possible downward shift in demand because of lower reliability of domestic honey supplies when much of it is delivered to the CCC. Hence, the demand for U.S.-produced raw honey (H_P) may be expressed as

$$(12) \quad H_P = D(PP^e, P_H, P_I, P_O, X),$$

where PP^e is the expected price for processed honey, P_H is the price of U.S. raw honey, P_I is the price of imported honey, P_O is a vector of prices of other inputs and X is defined above.

Processor demand for imported honey is expressed similarly to the U.S. demand:

$$(13) \quad H_I = I(PP^e, P_H, P_I, P_O, X).$$

The processor expectation model is discussed with the empirical specifications.

Farm-Level Demand for Beeswax

Beeswax is purchased by bee supply dealers who use the wax in the production of frame foundations and by manufacturing industries for cosmetics, pharmaceuticals, sculpturing, and candles. Imported wax is a strong substitute for domestically produced wax. Hence, the derived demand facing beekeepers expresses the price (P_W) as a function of the quantity of wax (W), the price of imported wax (P_{WI}), and other factors (Y) which may shift demand as seen by

$$(14) \quad P_W = W(W, P_{WI}, Y, X).$$

The variable X allows for a possible shift in wax demand among beekeepers under an effective honey price support program. The price of wax and the quantity demanded of wax are jointly determined. However, the demand equation is expressed as a price-dependent function because the quantity of wax is largely a predetermined variable.

Demand for Pollination Services

The demand for pollination services is derived from the input demand functions of farmers who produce crops requiring these services. This relationship is expressed as

$$(15) \quad S = S^D(P_S, P_F^e, P_Z),$$

where P_S is the price of pollination services, P_F^e is a vector of the expected prices of the products obtained from the services, and P_Z is a vector of prices of other farm inputs.

Demand and Market Allocation of Processed Honey

The annual supply of processed honey consists of the quantity of domestically produced raw

honey (H_P), which is converted to processed honey with little or no loss, the quantity of honey imported (H_I), and inventory carried into the current year (ST). Processor decisions on the amount of the annual supply to market in the current year (H_M) and to carry as inventory to the next year (ST_{t+1}) are made under conditions of uncertainty as to supply and demand conditions applicable to ST_{t+1} . There are no generally accepted theoretical principles to predict processor behavior in the face of this uncertainty. The formulation here is a modification of a model for canned fruit processors developed by French and King. Noting that processors have an opportunity to adjust outputs in $t + 1$, French and King argue that processors attempt to achieve an allocation in t to assure that the price covers the raw product and processing cost. However, they will be influenced further by their price and cost expectations in $t + 1$. The deflated lagged price and cost reflect the most recent market and cost experience and serve as a base for projecting possible sales in year $t + 1$.

With these considerations, a market allocation function may be approximated by

$$(16) \quad H_M = M(H_P + H_I + ST, PP, PP_{t-1}, PC, PC_{t-1}, X),$$

where H_M is quantity of honey marketed in t , H_P is quantity of U.S. honey processed, H_I is imported honey, ST is carry-in stocks, PP is processed honey price, PC is the raw product and processing cost. The variable X is included to test for a possible shift in market allocation decisions under an effective price support program.

Honey processors face a demand function derived from consumer, institutional, and manufacturing uses. With equation (16) normalized on quantity, the demand function is normalized on price as seen by

$$(17) \quad PP = P(H_M, N, X),$$

where N represents other variables (to be defined in the empirical specification) that account for shifts in the level of consumer and intermediary demand. The variable X accounts for possible effects of free distribution of CCC stocks beginning with their accumulations in 1981. Statistics on actual CCC distributions are not available.

Carry-over stocks are determined by the identity

$$(18) \quad ST_{t+1} = H_P + H_I + ST - H_M.$$

Equations (16), (17), and (18) form a simultaneous system.

Total System

With added specifications for how price expectations are formed, plus required identities, and the simplifying assumption that the supply of imported honey (H_I) is nearly perfectly elastic, equations (3) to (18) form a complete system that may be solved for the annual outputs and prices of all honey products, given the values of the exogenous variables and any lagged endogenous variables.

Empirical Model

The equations of the model are specified as linear approximations in the parameters but with some nonlinear identities. The variable X could affect all of the parameters of the equations in which it is included. However, because the observations to test this are insufficient, it is entered as a linear variable affecting only the level of response. Prices are deflated by the personal consumption expenditure deflator. All observations are annual values. The empirically estimated model is presented in table 1 with t -statistics in parenthesis below each coefficient and the equations' Durbin-Watson or Durbin h -statistic presented. The variables are defined in table 2. Estimation procedures are described following the discussion of empirical specifications.

Investment in Colonies

The colony response function, equation (3), is modified for empirical analysis by replacing the input prices with measures of unit production cost and combining the costs and product prices into measures of joint profitability (ratios of revenue per colony to cost per colony) for beekeepers specializing in honey production ($FHOPMT$), pollination services ($FPOPMT$), and bee production ($FPKPMT$) [see table 1, equations (3.2), (3.3), and (3.4) for definitions.] The number of colonies is expressed as a function of expected future values of the profitability measures.

In forming their future profitability expectations, beekeepers are assumed to rely mainly on average profit experience over recent periods.

Table 1. Econometric Model of the U.S. Honey Industry**Colony Response****Colony:**

$$(3.0) \quad COL_{it} = 139.658 + 0.903 COL_{it-1} + 242.299 FACMT_{it} \quad (\text{Durbin } h = 0.186)$$

(1.102) (34.323) (3.591)

Average profitability (lagged endogenous):

$$(3.1) \quad FACMT_{it} = (1/3) * [(FHOPMT_{it-1} + FHOPMT_{it-2})/2 + (FPOPMT_{it-1} + FPOPMT_{it-2})/2 + (FPKPM_{it-1} + FPKPM_{it-2})/2]$$

Honey profitability:

$$(3.2) \quad FHOPMT_{it} = [PHMAXD_{it} * WHOHO + PWXD_{it} * WWXHO + PPOD_{Cit} * WPOHO + PPKD_{Cit} * WPKHO + PQND_{Cit} * WQNH0]/(PPKD_{Cit} * QPKFO + PQND_{Cit} * QQNH0 + CHOPXD_{it})$$

Bee production profitability:

$$(3.3) \quad FPKPM_{it} = [PHMAXD_{it} * WHOPK + PWXD_{it} * WWXPK + PPOD_{Cit} * WFOPK + PPKD_{Cit} * WPKPK + PQND_{Cit} * WQNPK]/(PPKD_{Cit} * QPKFO + PQND_{Cit} * QQNPK + CPKPKD_{it})$$

Pollination profitability:

$$(3.4) \quad FPOPMT_{it} = [PHMAXD_{it} * WHOPO + PWXD_{it} * WWXPO + PPOD_{Cit} * WPOPO + PPKD_{Cit} * WPKPO + PQND_{Cit} * WQNPO]/(PPKD_{Cit} * QPKFO + PQND_{Cit} * QQNPO + CPOPXD_{it})$$

Farm price maximum:

$$(3.5) \quad PHMAXD_{it} = \text{MAXIMUM}(PHFD_{it}, PHSD_{it})$$

Product Supply and Demand**Honey supply:**

$$(4.0) \quad QHF_{it} = 121.935 + 0.049 COL_{it} + 117.232 FHOPMT_{it} - 73.478 FPKPM_{it} - 230.157 FPOPMT_{it} - 867.204 X_{it} \quad (D-W = 2.330)$$

(3.347) (5.749) (3.989) (-3.687) (-3.571) (-2.079)

Wax supply:

$$(5.0) \quad QWX_{it} = WXHOR_{it} * QHF_{it}$$

Pollination price setting:

$$(6.0) \quad PPOD_{Cit} = 11.063 + 0.423 PPOD_{Cit-1} + 0.0044 QPO_{Cit} - 0.0015 COL_{it-1} + 3.710 PHMAXD_{it-1} - C.191 TRND_{it} \quad (\text{Durbin } h = 0.414)$$

(4.300) (4.245) (5.377) (-3.471) (3.465) (-4.341)

Package price setting:

$$(7.0) \quad PPKD_{Cit} = 0.194 + 9.442 PHFD_{it-1} \quad (D-W = 1.482)$$

(1.639) (19.302)

Queen price setting:

$$(8.0) \quad PQND_{Cit} = -0.229 + 0.865 PPKD_{Cit} + 3.045 QQNCOL_{it} \quad (D-W = 1.469)$$

(-2.912) (24.562) (4.326)

Package bee demand:

$$(9.0) \quad QPKCOL_{it} = 0.035 - 0.026 PPKD_{Cit} + 0.243 PHMAXD_{it-1} + 0.933 QQNCOL_{it} - 0.894 X_{it} + 0.029 DUM65_{it} \quad (D-W = 1.770)$$

(5.788) (-3.499) (2.944) (10.942) (-3.954) (6.038)

Queen demand:

$$(10.0) \quad QQNCOL_{it} = -0.113 - 0.022 PQND_{Cit} + 0.289 PHMAXD_{it-1} + 0.247 QPKCOL_{it} - 0.169 X_{it} + 0.0023 TRND_{it} \quad (D-W = 1.414)$$

(-2.602) (-3.577) (4.489) (4.205) (-1.023) (6.200)

Allocation of honey between CCC and processors:^a

$$(11.0) \quad QHC_{it} = AHC_{it} * QHF_{it}$$

$$(11.1) \quad AHC_{it} = \text{POS}(-1.217 + 1.441 PHSFARD_{it}) \quad (D-W = 2.117)$$

(-3.059) (4.131)

$$(11.2) \quad PHSFARD_{it} = PHSD_{it}/PHFD_{it}$$

$$(11.3) \quad QHP_{it} = (1 - AHC_{it}) * QHF_{it}$$

Table 1. Continued

Demand for beekeepers' honey:

$$(12.0) \quad PHFD_t = 0.263 - 0.0044 QSHPM_t - 0.0028 ICHPD_t + 0.249 PHRDF_{t-1} + 0.017 DHM_{t-1} \\ (4.796) \quad (-0.393) \quad (-5.714) \quad (3.939) \quad (1.205) \\ + 0.613 PHID_t + 0.098 DUM73_t - 0.623 X_t \quad (D-W = 1.526) \\ (13.629) \quad (8.580) \quad (-2.214)$$

$$(12.1) \quad QSHPM_t = QHP_t/M_t + SHP_t/M_t$$

Demand for imported honey:

$$(13.0) \quad IHM_t = 0.375 - 0.143 QSHPM_t + 0.827 PHMAXD_t + 0.068 PHRDF_{t-1} - 1.699 PHID_t \\ (3.570) \quad (-4.735) \quad (1.558) \quad (0.282) \quad (-4.035) \\ + 0.143 DUM73_t + 2.667 X_t \quad (D-W = 1.848) \\ (2.562) \quad (2.465)$$

Wax demand:

$$(14.0) \quad PWXD_t = 0.151 - 5.713 QWXM_t + 0.055 FHOPMT_{t-1} + 0.786 PWXID_t - 2.159 X_t \quad (D-W = 1.500) \\ (3.447) \quad (-4.930) \quad (2.529) \quad (25.915) \quad (-3.648)$$

$$(14.1) \quad QWXM_t = QWX_t/M_t$$

Processors' Marketing

Domestic supply of processed honey:

$$(16.0) \quad QDHMM_t = -0.295 + 0.943 QSHPM_t + 1.179 PHRDFX_t - 0.250 PHMAXDX_t \\ (-2.287) \quad (15.472) \quad (2.276) \quad (-0.583) \\ + 0.0058 TRND_t - 5.049 X_t \quad (D-W = 1.586) \\ (2.385) \quad (-4.692)$$

$$(16.1) \quad PHRDFX_t = PHRDF_t - PHRDF_{t-1}$$

$$(16.2) \quad PHMAXDX_t = PHMAXD_t - PHMAXD_{t-1}$$

Demand for processed honey:

$$(17.0) \quad PHRDF_t = 0.423 - 0.043 DHM_t + 0.213 DUM73_t - 0.012 TRND73_t \\ (17.504) \quad (-2.257) \quad (21.798) \quad (-9.338) \\ - 0.993 X_t \quad (D-W = 2.381) \\ (-2.340)$$

$$(17.1) \quad DHM_t = QDHMM_t + IHM_t - EH_t/M_t$$

Carry-over stocks:

$$(18.0) \quad SHP_{t+1} = QHP_t + (IHM_t * M_t) + SHP_t - (DHM_t * M_t) - EH_t$$

* In (11.1), the POS function takes the value in parentheses or 0, whichever is larger.

Profit measures averaged over various lags were tested, with a two-year average providing the best predictions of colony numbers. The lagged profit measures for the three beekeeper specializations were combined into a single measure of lagged average profitability (*FACMT2*) (see table 2, equation (3.1)). The variable *X* was not statistically significant in the empirical estimation and was omitted from the final equation.

Data pertaining to diseases, pesticide losses, and extreme weather which may affect colony numbers are incomplete or unavailable, so these effects are reflected as part of an unexplained disturbance.

Beekeepers' desired level of colonies may not be achieved instantaneously because of the time required to fill orders, make deliveries, and obtain capital. Actual and desired colonies are assumed to be related as in the partial adjustment

model. That is, actual colonies in *t* are equal to actual colonies in *t* - 1 plus a proportion of the difference between desired colonies in *t* and actual colonies in *t* - 1.

With these considerations, the colony response function is expressed linearly as equation (3.0) in table 1 where the variables are defined in table 2.

Honey and Wax Supply

As in the colony response function, the price variables in the conditional supply function for honey, equation (4), are converted to measures of unit costs and returns and combined into profitability measures for the three major specializations as seen in equation (4.0) in table 1. The supply of wax is proportional to honey pro-

Table 2. Empirical Model Variable Definitions

Name	Definition	Measure
<i>AHC</i>	Allocation of honey to the CCC	(proportion)
<i>CHOPXD</i>	Exogenous input costs for honey producer	(\$/colony)
<i>*COL_t</i>	Industry colonies	(thousands)
Constant	Intercept	(1)
<i>CPKPYD</i>	Exogenous input costs for package bee producer	(\$/colony)
<i>CPOPXD</i>	Exogenous input costs for pollination producer	(\$/colony)
<i>*DHM</i>	Disappearance of honey	(lbs/person)
<i>DUM65</i>	Dummy in 1965 and after	(0 or 1)
<i>DUM73</i>	Dummy in 1973 and after	(0 or 1)
<i>EH</i>	Exports of honey	(million lbs)
<i>*FACMT2</i>	Profitability ratio for all products in time $t - 1$ and $t - 2$	(dimensionless)
<i>*FHOPMT</i>	Profitability ratio for honey production	(dimensionless)
<i>*FPKPYD</i>	Profitability ratio for package bee production	(dimensionless)
<i>*FPOPMT</i>	Profitability ratio for pollination services	(dimensionless)
<i>ICHPD</i>	Index of costs of honey processing	(1972 = 100)
<i>IHM</i>	Imports of honey	(lbs/person)
<i>M</i>	Population	(millions)
<i>*PHFD</i>	Farm price of honey	(1972\$/lb)
<i>PHID</i>	Price of U.S. honey imports	(1972\$/lb)
<i>*PHMAXD</i>	Maximum farm price of honey	(1972\$/lb)
<i>PHMAXDX</i>	Farm price of honey differential between t and $t - 1$	(1972\$/lb)
<i>*PHRDF</i>	Retail price of honey	(1972\$/lb)
<i>PHRDFX</i>	Retail price of honey differential between t and $t - 1$	(1972\$/lb)
<i>PHSD</i>	Price support for honey	(1972\$/lb)
<i>PHSFARD</i>	Support to farm honey price ratio	(dimensionless)
<i>PPKD_c</i>	Price of package bees (California)	(1972\$/lb)
<i>*PPOD_c</i>	Price of pollination services (California)	(1972\$/service)
<i>PQND_c</i>	Price of queen bees (California)	(1972\$/bee)
<i>PWXD</i>	Price of wax	(1972\$/lb)
<i>PWXID</i>	Price of wax imports	(1972\$/lb)
<i>QDHMM</i>	Quantity of domestic honey marketed	(lbs/person)
<i>QHC</i>	Quantity of honey to the CCC	(millions lbs)
<i>QHF</i>	Quantity of honey	(million lbs)
<i>QHP</i>	Quantity of honey to processors	(million lbs)
<i>QPKCOL</i>	Ratio of packages to colonies in $t - 1$	(lbs/colony)
<i>**QPK(J)</i>	Packages used by (J) producer, where $J = HO$ Honey PK Package bee PO Pollination	(lbs/colony)
<i>QPO_c</i>	Quantity of pollination services (California)	(thsd services)
<i>QQNCOL</i>	Ratio of queens to colonies	(bees/colony)
<i>**QQN(J)</i>	Queens used by (J) producer, where J is as in <i>QPK(J)</i>	(bees/colony)
<i>QSHPM</i>	Total domestic quantity of honey at the processor	(lbs/person)
<i>QWX</i>	Quantity of wax	(million lbs)
<i>QWXM</i>	Quantity of wax	(lbs/person)
<i>*SHP</i>	Stocks of honey	(million lbs)
<i>TRND</i>	Linear time trend	(year, 1952 = 3)
<i>TRND73</i>	Time trend beginning in 1973	(year, 1973 = 1)
<i>**WHO(J)</i>	Honey produced by (J) producer, where J is as in <i>QPK(J)</i>	(lbs/colony)
<i>**WPK(J)</i>	Packages produced by (J) producer, where J is as in <i>QPK(J)</i>	(lbs/colony)
<i>**WPO(J)</i>	Pollination services produced by (J) producer, where J is as in <i>QPK(J)</i>	(services/colony)
<i>**WQN(J)</i>	Queens produced by (J) producer, where J is as in <i>QPK(J)</i>	(bees/colony)
<i>**WWX(J)</i>	Wax produced by (J) producer, where J is as in <i>QPK(J)</i>	(lbs/colony)
<i>WXHOR</i>	Wax to honey production ratio	(lbs/lbs)
<i>X</i>	Dummy variable for support program effectiveness	(0 or 1972\$/lb)

duction as seen in equation (5.0), where $WXHOR$ is an exogenous variable that specifies the ratio of wax production to honey production.

Supply and Demand for Pollination Services

Time-series data on the price and quantity of pollination services are available only for California (the major user of such services). Hence, the pollination component of the model refers only to California, although some pollination services in California may be obtained from beekeepers outside of the state.

The supply equation, (6), is determined jointly with considerations of the demand for pollination services, equation (15). Because there is little or no substitution between pollination and other inputs and because the cost of pollination relative to the total value of crop production is quite small, the demand for pollination services is very inelastic. The primary factors determining the demand for pollination services in California (QPO_C) are the area of land requiring pollination services and the number of pollination services used per acre. The latter increased over the period of study due in part to changes in the crop mix and farmers' greater awareness of the potential benefits of bees. Empirical analysis revealed no significant effect of the pollination price on the pollination services used per acre over the range of observed data. Hence, the demand for pollination services is expressed as an exogenous variable rather than as indicated by equation (15).

With the quantity of pollination services (QPO_C) entered exogenously, the supply function for pollination services contains price as the dependent variable. The price of pollination services ($PPOD_C$) is a function of the price of the service charged in the previous year, the quantity of services demanded (QPO_C), the availability of colonies to provide these services (COL_t), the price received to produce honey ($PHMAXD$), and a time trend ($TRND$) as seen in equation (6.0).

The previous period's pollination price reflects inertia in the system and existing contractual arrangements between farmers and beekeepers. The colonies available to provide the services are based on previous period values since current-year colony numbers are not known at the time pollination prices are set in very early spring. Lagged honey price reflects the trade-off between honey production and pollination services resulting from the foraging intensity of

providing pollination services and the bitter taste of honey from almond pollination. Variable $TRND$ accounts for secular increases in beekeepers' willingness to supply colonies for pollination services. Variable X , reflecting changes in the support program, was not significant in the empirical estimation.

Supply and Demand for Package Bees and Queens

As in the case of pollination services, price and quantity time-series data for package bees and queen bees are available only for California. Therefore, the package bee and queen bee supply relationships refer to California.

Discussions with industry representatives suggested that package bee producers, who have the option of retaining the bees for their own use, base the package price on the price of honey and sell whatever packages are demanded at that price. The price equation is specified as equation (7.0) in table 1. Because the current-year average honey price is not known at the time most packages are sold, last year's price is used in the specification. This price-setting equation replaces the supply equation (7) in the system.

The price of queen bees has been set based on the price of packages, with some modification reflecting the movement of queens relative to the number of colonies. The relationship is expressed as equation (8.0) in table 1.

The demands for package bees and queens are proportional to the number of colonies. However, the proportions vary with the product price, the level of the federal support program and shifts over time that are difficult to measure. The two demand functions are expressed as equations (9.0) and (10.0) in table 1. The ratio of packages to colonies ($QPKCOL$) is in equation (10.0) and the ratio of queens to colonies ($QQNCOL$) is in equation (9.0) to account for the complementarity in the two demands.

A shift in demand for packages occurred in the mid-1960s. The exact cause of the change is unknown, but it could be attributable to increased bee kills from new chemicals and the need for replacement bees or the increased acceptance of starting colonies from packages and shipping these packages through the mail system. The variable $DUM65$, with a value of zero prior to 1965 and one thereafter, was introduced in the package demand equation (9.0) to account for the change.

The demand for queens did not shift signifi-

cantly in 1965 as did the demand for packages. Queens are demanded by beekeepers for a more specialized purpose than packages. As the science surrounding artificial insemination of queens and the isolation of genetic qualities has improved, the increased confidence beekeepers place in queen bees is manifested by increased demand for queens. The effect of this scientific evolution on the demand for queen bees is accounted for by a linear trend variable (*TRND*).

Allocation of Honey between Processing and the CCC

The empirical form of equation (11) is specified by equations (11.0), (11.1), (11.2), and (11.3) in table 1. The *POS* function indicates *AHC* takes the value in parenthesis or zero, whichever is larger.

Processor Demand for Honey

The empirical form of the processor demand function for U.S.-produced honey, equation (12), replaces the vector of processing input prices (P_o) with an index of processing cost (*ICHPD*) and normalizes the equation on price, rather than quantity. The price-dependent form is selected because honey supplies are partially predetermined and the intercorrelation problem among prices is reduced. All quantities are expressed on a U.S. per capita basis. The price processors expect to receive for the processed product (PP^e) is affected by carry-in stocks and imports and perceived shifts in the level of per capita demand. The latter may be indicated by changes in lagged deflated processed product prices and per capita movement. Past period price changes while movement is constant indicate a shift in demand; similarly, increased movement while price remains constant would indicate a shift in demand.

An upward shift in the demand for honey also occurred in 1973 that was not accounted for by the lagged price and movement variables. One possible explanation is a change in the buyer psychology during and following the energy crisis of 1973 and 1974. The variable *DUM73* accounted for this shift. With these considerations, the demand function facing beekeepers is specified as equation (12.0) in table 1.

The demand for imported honey [equation (13)] is a function of the import price (*PHID*), the domestic honey price (maximum of support and

market price *PHMAXD*), and the price at which processors expect to sell the honey. Substituting for expected price as in equation (12) and simplifying, the demand for imported honey can be expressed as equation (13.0) in table 1. The lagged per capita movement of honey (DHM_{t-1}) was not significant in the empirical estimation and was omitted from the final equation. The variable *DUM73* also was introduced in this relationship.

Because of limited data the impacts of foreign supply and demand on U.S. imports could not be evaluated. No statistically significant relationship was found between U.S. import honey demand and exchange rates. To simplify the analysis, import supply is assumed to be approximately perfectly elastic at given prices.

Farm-Level Demand for Beeswax

Equation (14) in the conceptual model is represented by equation (14.0) in table 1. The variable *QWXM* replaces *W*, *PWXID* replaces P_w and *FHOPMT* replaces *Y*. *FHOPMT* accounts for changes in the wax demand from beekeepers. As the profitability of the honey industry increases one would anticipate expansion in the number of colonies and a greater demand for wax frame foundations.

Demand Facing Honey Processors and Their Market Allocation

The market allocation for honey, equation (16), is specified as equation (16.0) in table 1. The quantities imported (*IHM*) are sold currently with no inventory carryover because imports likely are coordinated closely with market demands. The dominant factor affecting the quantity of domestically produced honey marketed is the available supply of domestic honey (pack plus carry-in stocks). An increase in costs from the previous period may lead processors to reduce shipments (and increase carryover) in order to raise current price. If the current price increases relative to the previous period price, current shipments may be increased; if price decreases, shipments may be reduced with the hope that prices will be improved the next period. The change in *ICHPD* was not statistically significant in the empirical estimation and was omitted from the final equation. The *TRND* variable accounts for a general increase in market allocation and reduced average inventory carryover.

The domestic demand facing honey processors is expressed as equation (17.0) in table 1 where *DHM* is defined as equation (17.1).² The data suggest that the per capita demand for honey remained stable until about 1973. It then shifted upward with the onset of high-level inflation and the Arab oil embargo but with a down trend because of increasing health and caloric concerns about sweeteners. The dummy shift variable, *DUM73*, and the trend variable beginning in 1973, *TRND73*, accounted for these complex shifts. *DUM73* has a value of zero prior to 1973 and is 1 thereafter. *TRND73* is a trend variable that is zero prior to 1973, then is 1 in 1973, 2 in 1974, etc.

Because data for the honey industry are available only for calendar years rather than seasonal years, stocks are measured as of January 1. The major domestic honey flow is in the spring and summer. With seasonal year data, April 1 to April 1, the seasonal supply to be allocated is predetermined for a given year. However, with calendar year data, the quantity packed (and therefore the annual supply) is determined simultaneously with the processors' allocation of the processed product, the price of the processed product, the carryover stock level and the demand for imported honey.

In terms of the empirical variable definitions, equation (18) becomes equation (18.0) in table 1, where *SHP* refers to January 1 stocks. Other variables are as defined in table 2.

Empirical Results

The complete model consists of thirteen jointly related stochastic equations, linear in their parameters, and fifteen identities or technical equations seen in table 1. One of the technical equations, equation (3.1), describes *FACMT2*, a calculated predetermined variable rather than an endogenous variable. Some identities, such as equations (3.2), (3.3) and (3.4), are nonlinear. There are twenty-seven current endogenous variables, eleven lagged endogenous variables, and seventeen exogenous variables in the model. The exogenous variables are underlined in the variable definitions (table 2). Lagged endoge-

nous variables are preceded by an asterisk. Parameters indicating the proportion of each product produced by beekeepers specializing in honey production, bee production, or pollination services and parameters indicating the quantity of packages and queens purchased by each specialty are derived from Reed and Horel. They are identified by a double asterisk.

The data are calendar year values for the period 1952 to 1984. Data for 1985, 1986, and 1987 are used for out-of-sample prediction tests. Data pertaining to bee colonies, honey quantities, and prices are U.S. values. Data pertaining to cost of production, prices and quantities of pollination services, packaged bees and queen bees are for California since U.S. values are not reported. Monetary values are deflated by the U.S. personal consumption expenditure deflator. Data pertaining to prices received by honey processors were available for only part of the period of analysis. Therefore, a somewhat longer series of retail prices was used as a proxy for the wholesale prices. Reporting of the retail price series was discontinued after 1979. Hence, the prices for 1980 to 1984 were estimated from historical retail-farm price margins (Willett, pp. 135-37).

The model was estimated by three-stage least squares (3SLS) except for equation (11.1), which involves a limited dependent variable. The existence of a limited dependent variable within a set of simultaneous equations involves complex estimation procedures and restrictions (Amemiya). To simplify the analysis, this equation was estimated by ordinary least squares with data for 1981 to 1984 because this is the period the support program essentially was effective. Separate estimation of equation (11.1) may affect the efficiency of the estimates, but the loss likely is small compared with the gain in estimation simplicity.³

Performance Characteristics

All estimated coefficients have signs consistent with the theoretical specifications of the model. Most are large relative to their standard errors, as indicated by the *t*-statistics in parentheses. The Durbin-Watson or Durbin *h*-statistics for all equations either do not reveal autoregressive re-

² Other sweeteners, such as sugar, corn syrup, and artificial sweeteners, are not considered significant honey substitutes because of honey's unique characteristics (color, flavor, viscosity, moisture retention capabilities, and marketing appeal). Recent promotional efforts by the National Honey Board have increased the attractiveness of honey and its use in food products.

³ Estimation of equation (11.1) by maximum likelihood Tobit for the period 1952-84 yields similar results. The ordinary least squares estimates based on 1981-84 data were used since they provide a closer fit for the period of policy analysis.

ror structures at the 5% significance level or are inconclusive.

The ability of the model to provide reasonably close fits to the historical data used in the estimation was evaluated by static one-period-ahead predictions of the model's endogenous variables. The predictions were made given the values of the exogenous variables, the carry-in stocks of honey and actual values of other predetermined endogenous variables. Most average absolute prediction errors were within 10% of the mean values of the predicted variables.

The stability of the nonlinear model was tested by holding the exogenous variables constant (at their 1984 values) and allowing the model to generate values of the endogenous variables over a future period. All variables approached stationary equilibrium values by the fifteenth time period, suggesting the model is stable.

To test for possible structural changes or specification errors, out-of-sample predictions were calculated for the period 1985, 1986, and 1987, subject to the limitation of available data. The model's predictions were obtained by inserting observed values of right-side variables in each structural equation. When data were not available for some exogenous variables, they were assumed to remain at 1984 values. Actual colony levels for 1986 and 1987 were assumed to remain at 1985 levels because the colonies reported by the U.S. Department of Agriculture (USDA) decreased from 4,325,000 in 1985 to 3,205,000 and 3,190,000 in 1986 and 1987, respectively. This unprecedented drop of 1 million colonies results from USDA's elimination of reporting those colonies of beekeepers with fewer than five hives.

In 1986, a lower loan repayment option allowed producers to repay their loans at a rate below the support rate. Hence, producers were encouraged to repay their loans and sell their honey on the market. Because of the change in the federal honey program in 1986, calculation of the variable reflecting adjustments to the honey support program (X) was no longer consistent with previous years. However, if this variable is set to zero, the model overpredicts several endogenous variables. Because of the continuing uncertainty about the existence of the honey support program, it seems reasonable to hold the variable X at its 1985 value for testing purposes.

The model's predictions, seen in table 3, are reasonably close to actual values and are within the 99% confidence interval suggested by the standard error of the regressions (identified by the S under each equation) for all but three vari-

ables.⁴ The model underpredicts the production of honey (Q_{HF}) in 1987 by an amount inconsistent with the values of the standard error. The 1987 honey crop was large because of excellent weather conditions across the United States. In addition, it appears that by 1987 the depressing effect of the uncertainty surrounding the support program was no longer as strong as it was in 1985 and 1936. Elimination of the effect of the support program in 1987 would mean an increase in the model's prediction of honey production by 47 million pounds. With this alteration the actual level of production would be within the 99% confidence interval suggested by the standard error of the regression.

The 1987 prediction of per capita honey imports (I_{HM}) is also outside the confidence interval. Thus, the federal honey support program did not appear to influence the quantity of honey imported in 1987 as the model suggests. Elimination of this effect (setting $X = 0$) means the model's predictions of per capita imports for 1987 would be 0.250 pounds per person, clearly within the range of the confidence interval.

Finally, the predictions of the per capita quantity of domestic honey marketed (Q_{DHMM}) were out of bounds in 1986. Processors did not increase their marketing of domestic honey as quickly as the model predicts. The model predicts processors would market more of their honey and hold less in stocks than actually occurred. By 1987 the model's predictions were closer to actual values.

Interpretation and Policy Analysis

Flexibilities and elasticities, evaluated at the mean of the data set and at 1980 values, for the model's supply and demand relationships are presented in table 4. The values for 1980 were chosen because that is the most recent year the support program was not effective. The variable $\epsilon_{COL1,PHMAXD}$ refers to colony response to the farm price of honey while $\epsilon_{COL1,FACMT2}$ refers to colony response to average profitability of all beekeeping products. The variable $\epsilon_{QHF,PHMAXD}$ is the conditional supply elasticity for honey given the

⁴ Model predictions might be examined in relation to confidence intervals based on the equation's standard error of the forecast rather than the standard error of the regression. In general, the standard forecast error will not be less than the standard error of the regression. Because the standard errors of the forecast involve complex calculations and the predictions for most variables fall within the reported confidence intervals, these standard errors were not calculated.

Table 3. Comparison of Actual and Predicted Values of Endogenous Variables in the Honey Industry Model 1985-87

	Year	Actual (A)	Predicted (P)	Difference (A-P)
Colony, COL_t	1985	4325.0	4266.1	58.9
Equation (3.0)	1986	4325.0	4303.1	21.9
$S = 74.888$	1987	4325.0	4306.9	18.1
Honey supply, QHF	1985	150.1	152.3	-2.2
Equation (4.0)	1986	200.4	151.3	49.1
$S = 20.055$	1987	226.8	145.5	81.3
Pollination price, $PPOD_c$	1985	8.676	8.683	-0.007
Equation (6.0)	1986	8.292	8.461	-0.169
$S = 0.314$	1987	8.183	8.072	0.110
Package price, $PPKD_c$	1985	2.554	2.233	0.320
Equation (7.0)	1986	2.057	2.068	-0.010
$S = 0.214$	1987	1.970	2.173	-0.203
Queen price, $PQND_c$	1985	2.151	2.319	-0.168
Equation (8.0)	1986	2.145	1.890	0.255
$S = 0.144$	1987	2.016	1.842	0.174
Package demand, $QPKCOL$	1985	0.119	0.120	-0.001
Equation (9.0)	1986	0.114	0.130	-0.016
$S = 0.011$	1987	0.128	0.125	0.003
Queen demand, $QQNCOL$	1985	0.111	0.118	-0.007
Equation (10.0)	1986	0.111	0.116	-0.005
$S = 0.008$	1987	0.121	0.110	0.011
Allocation, AHC	1985	0.653	0.764	-0.111
Equation (11.0)	1986	0.208	0.224	-0.016
$S = 0.080$	1987	0.208	0.224	-0.017
Honey price, $PHFD$	1985	0.198	0.189	0.009
Equation (12.0)	1986	0.210	0.189	0.020
$S = 0.011$	1987	0.197	0.190	0.007
Import demand, IHM	1985	0.578	0.606	-0.029
Equation (13.0)	1986	0.490	0.454	0.036
$S = 0.042$	1987	0.239	0.401	-0.162
Wax price, $PWXD$	1985	0.484	0.461	0.023
Equation (14.0)	1986	0.383	0.439	-0.056
$S = 0.028$	1987	0.420	0.416	0.004
Honey marketed, $QDHMM$	1985	0.082	0.152	-0.070
Equation (16.0)	1986	0.349	0.744	-0.396
$S = 0.052$	1987	0.999	1.067	-0.068
Honey processing price, $PHRDF$	1985	0.403	0.398	0.005
Equation (17.0)	1986	0.415	0.379	0.036
$S = 0.017$	1987	0.402	0.350	0.052

level of colonies and $\epsilon_{QDHMM,PHRDF}$ is the allocation elasticity for honey given the seasonal supply. The variables f_{PPOD_c,QPO_c} and f_{PQND_c,QQN_c} are the price flexibilities for pollination and queen pricing, respectively. Package price, equation (7.0), is based on the honey price and has no flexibility. The demand elasticity for imported honey ($\epsilon_{IHM,PHID}$) is elastic while the demand elasticities for packages ($\epsilon_{QPKCOL,PPKD_c}$) and queens ($\epsilon_{QQNCOL,PQND_c}$) are inelastic. The flexibilities for wax price ($f_{PWXD,QWXM}$), domestic honey producer price ($f_{PHFD,QHFM}$) and processor price ($f_{PHRDF,DHM}$) are low.

An important problem in estimating the model was accounting for difficult-to-measure shifts in levels of some of the model equations, partly

because of honey price supports and partly because of other structural shifts. With regard to the latter, the price of pollination services, equation (6.0), decreased over the period of the analysis at the rate of 0.191 cents per year. The data suggest that package bee demand, equation (9.0), shifted upward in 1965 although the reasons are not entirely clear. The demand for queens, equation (10.0), has shown a general upward trend due to changes in production practices. The demands for honey, equations (12.0), (13.0), and (17.0), shifted in a complex manner. The demand for processed honey shifted upward about 1973 (consistent with the general structural change for a number of agricultural commodities), but demand trended down fol-

Table 4. Price Elasticities and Flexibilities

Elasticity/ Flexibility		Evaluated at the Mean	Evaluated at 1980 Values
Colony			
$\epsilon_{COL,PHMAXD}$	(short run)	0.024	0.023
	(long run)	0.242	0.237
$\epsilon_{COL,FACMT2}$	(short run)	0.059	0.067
	(long run)	0.610	0.687
Supply			
$\epsilon_{QHF,PHMAXD}$		0.193	0.182
$f_{PPODC,QPOC}$	(short run)	0.410	0.621
	(long run)	0.710	1.182
$f_{QNDQ,QQNC}$		0.111	0.182
$\epsilon_{QDHMM,PHRDF}$		0.447	0.575
Demand			
$f_{FWXD,QWXM}$		-0.164	-0.099
$\epsilon_{QPKCOL,QPKDC}$		-0.517	-0.380
$\epsilon_{QQNCOL,QQNC}$		-0.580	-0.368
$f_{PHFD,QHPM}$		-0.021	-0.012
$\epsilon_{IHM,PHID}$		-2.246	-1.831
$f_{PHRDF,DHM}$		-0.121	-0.085

lowing 1973. The negative trend effects in the demand for processed honey likely are absorbed in the coefficients for *ICHPD* and *PHID* in the processor raw honey and import demand equations, equations (12.0) and (13.0), respectively.

The significance of variable *X* in several equations suggests that when the price support for honey exceeded the market price it had a significant impact on several sectors of the honey industry. No direct effect on the number of colonies maintained by beekeepers was evident; but, when the support program was effective, a decrease in beekeepers' supply of honey was reflected by a shift in beekeepers' specializations among honey, pollination, and packages. The demand for packages and queens all appeared to decrease when the support program was effective. The demand for wax and domestic honey decreased as imported products became more attractive. In addition, the demand for processed honey declined due to free CCC disbursements. Apparently, consumers substituted CCC disbursements for market priced honey. The demand for imported honey increased as processors chose not to rely on the uncertain domestic supply. The available supply of processed honey marketed in the current year fell because of the increased uncertainty about market conditions. Hence, the model shows that the federal honey support program caused a substitution of imported honey for domestic honey and shifted large quantities of honey to the CCC.

Table 5 captures actual and predicted values of the key endogenous variables of the model for the period 1982–85 assuming the price support had continued below the market price, i.e., $X = 0$. The years 1986–87 were omitted because of the change of the federal support program. The predictions were made with the estimated disturbance terms entered as exogenous variables. It seems reasonable that the random factors captured by the disturbance would not change under alternative scenarios. Since adding the disturbance term gives the exact prediction for the historical model when *X* is included, the model predictions with *X* equal zero are compared with actual values rather than deterministic predictions.

Without an effective support program the market price for honey (*PHFD*) increased. Yet, this price is still less than the support price beekeepers were receiving under the federal program. In response to the price change, the level of colonies (*CCL_t*) decreases slightly and the price of pollination services (*PPOD_c*) falls. The prices of package bees (*PPKD_c*) and queen bees (*PQND_c*) increase, eventually increasing the profitability of these beekeeping specializations. In response to higher prices and profitability and less uncertainty about the support program, beekeepers use their colonies more intensely and increase their honey production (*QHF*) and the quantity of queens and package bees produced per colony (*QQNCOL*, *QPKCOL*). More domestic honey is sent to the processor and the quantity of honey imported (*IHM*) decreases. As less honey is forfeited to the CCC, free disbursements are not made. The demand for processed honey shifts and the price of honey received by processors (*PHRDF*) increases. Processors market a larger quantity of domestic honey (*QDHMM*).

Without the support program the predicted revenue received by beekeepers for their five products decreased slightly during 1982 through 1984 as seen in table 5. However, in 1985 beekeeper revenue is \$5 million more without an effective support program. Free disbursements by the CCC decreases. Hence, consumer expenditures for processed honey increase because of higher prices and increased per capita consumption of domestic honey. When the support program is not effective, the federal government has a potential savings of \$246 million from 1982 through 1985 (Hoff and Phillips). A comparison of the beekeeper revenue, higher consumer expenditures, and savings to the federal govern-

Table 5. Comparison of Actual and Predicted Values (Including Disturbances) of Endogenous Variables with and without the Support Program 1982-85

	Year	Support (S)	No Support (NS)	Difference (S-NS)
Colony, COL_t Equation (3.0)	1982	4250.0	4250.0	0.0
	1983	4275.0	4272.0	3.0
	1984	4300.0	4291.0	9.0
	1985	4325.0	4307.8	17.2
Honey supply, QHF Equation (4.0)	1982	230.0	230.4	-0.4
	1983	205.0	217.0	-12.0
	1984	165.1	187.8	-22.7
	1985	150.1	186.4	-36.3
Pollination price, $PPOD_c$ Equation (6.0)	1982	9.285	9.285	0.000
	1983	8.843	8.774	0.069
	1984	8.683	8.592	0.091
	1985	8.686	8.507	0.179
Package price, $PPKD_c$ Equation (7.0)	1982	2.929	2.929	0.000
	1983	2.580	2.587	-0.007
	1984	2.404	2.505	-0.101
	1985	2.554	2.790	-0.236
Queen price, $PQND_c$ Equation (8.0)	1982	2.499	2.504	-0.005
	1983	2.414	2.423	-0.009
	1984	2.311	2.421	-0.110
	1985	2.151	2.370	-0.219
Package demand, $QPKCOL$ Equation (9.0)	1982	0.172	0.177	-0.005
	1983	0.169	0.179	-0.010
	1984	0.145	0.175	-0.030
	1985	0.119	0.154	-0.035
Queen demand, $QQNCOL$ Equation (10.0)	1982	0.132	0.134	-0.002
	1983	0.128	0.129	-0.001
	1984	0.123	0.131	-0.008
	1985	0.111	0.116	-0.005
Allocation, AHC Equation (11.0)	1982	0.324	0.000	0.324
	1983	0.519	0.000	0.519
	1984	0.641	0.000	0.641
	1985	0.653	0.000	0.653
Honey price, $PHFD$ Equation (12.0)	1982	0.265	0.266	-0.001
	1983	0.244	0.255	-0.011
	1984	0.216	0.241	-0.025
	1985	0.198	0.240	-0.042
Import demand, IHM Equation (13.0)	1982	0.396	0.331	0.065
	1983	0.468	0.356	0.112
	1984	0.543	0.348	0.195
	1985	0.578	0.337	0.241
Wax price, $PWXD$ Equation (14.0)	1982	0.772	0.779	-0.007
	1983	0.605	0.632	-0.027
	1984	0.563	0.620	-0.057
	1985	0.484	0.568	-0.084
Honey marketed, $QDHMM$ Equation (16.0)	1982	0.568	0.853	-0.285
	1983	0.435	0.878	-0.443
	1984	0.138	0.804	-0.666
	1985	0.082	0.836	-0.754
Honey processing price, $PHRDF$ Equation (17.0)	1982	0.470	0.464	0.006
	1983	0.449	0.451	-0.002
	1984	0.421	0.434	-0.013
	1985	0.403	0.433	-0.030
Beekeeper revenue Million \$	1982	83.8	80.3	3.5
	1983	74.3	72.7	1.7
	1984	63.9	63.1	0.8
	1985	57.8	62.9	-5.2
Government payments Million \$	1982	27.4	0.0	27.4
	1983	48.0	0.0	48.0
	1984	90.2	0.0	90.2
	1985	80.8	0.0	80.8

ment suggests the honey support program was an ineffective means of supporting honey prices during the period of analysis.

[Received February 1988; final revision received March 1990.]

References

- Amemiya, T. "Multivariate Regression and Simultaneous Equation Models When the Dependent Variables Are Truncated Normal." *Econometrica* 42(1974):999-1012.
- Brandt, J. A., and B. C. French. *An Analysis of Economic Relationships and Projected Adjustment in the U.S. Processing Tomato Industry*. Berkeley: Giannini Foundation of Agr. Econ. Res. Rep. No. 331, 1981.
- Cheung, S. N. S. "The Fable of the Bees: An Economic Investigation." *J. Law and Econ.* 16(1973):11-33.
- French, B. C., and G. A. King. "Demand and Price Markup Functions for Canned Cling Peaches and Fruit Cocktail." *West. J. Agr. Econ.* 11(1986):8-18.
- Gould, J. R. "Meade on External Economies: Should the Beneficiaries Be Taxed?" *J. Law and Econ.* 16(1973):53-66.
- Hendry, D. F., and J. F. Richard. "On the Formulation of Empirical Models in Dynamic Econometrics." *J. Econometrics* 20(1982):3-33.
- Hoff, F. L., and J. K. Phillips. *Honey Background for 1990 Farm Legislation*. Washington DC: U.S. Department of Agriculture, Econ. Res. Serv. CED Staff Rep. No. AGES 89-43, 1989.
- Intriligator, M. D. *Econometric Models, Techniques, and Applications*. Englewood Cliffs NJ: Prentice-Hall, 1978.
- Johnson, D. B. "Meade, Bees and Externalities." *J. Law and Econ.* 16(1973):35-52.
- Meade, J. E. "External Economies and Diseconomies in a Competitive Situation." *Econ. J.* 62(1952):54-67.
- McDowell, R. *The Africanized Honey Bee in the United States: What Will Happen to the U.S. Beekeeping Industry?* Washington DC: U.S. Department of Agriculture, Econ. Res. Serv. NRED Agr. Econ. Rep. No. 519, 1984.
- Olmstead, A. L., and D. B. Wooten. "Bee Pollination and Productivity Growth: The Case of Alfalfa." *Amer. J. of Agr. Econ.* 69(1987):56-63.
- Reed, A. D., and L. A. Horel. *Bee Industry Economic Analysis for California*. Div. Agr. Sci. Leaflet No. 2345, University of California, 1976.
- Robinson, W. S., R. Nowogrodzki, and R. A. Morse. "The Value of Honey Bees as Pollinators of U.S. Crops." *Amer. Bee J.* 129(1989):411-23, 477-87.
- Siebert, J. W. "Beekeeping, Pollination, and Externalities in California Agriculture." *Amer. J. Agr. Econ.* 62(1980):165-71.
- U.S. General Accounting Office. *Federal Price Support for Honey Should Be Phased Out*. GAO/RCED-85-107, 1985.
- Willett, L. S. "An Econometric Analysis of the Supply and Demand Relationships in the U.S. Honey Industry." University of California, Davis, 1987.

Put-Call Parity and Arbitrage Bounds for Options on Grain Futures

William W. Wilson and Hung-Gay Fung

Many arbitrage opportunities have become available for market participants since the inception of trading in options. Central to many of these is the put-call parity relationship. If this condition is violated, arbitrage profits could be earned. This paper examines arbitrage possibilities between put and call options on grain futures using two different tests. One is a test of market efficiency over the duration of trading, and the other evaluates arbitrage bounds on individual trading days. In general, the results suggest that the options markets on grain futures are not completely efficient. However, violations varied substantially across contracts and through time.

Key words: arbitrage, market efficiency, options, put-call parity.

Since the inception of options trading on agricultural futures contracts in late 1984 and early 1985, many speculative, hedging, and arbitrage opportunities have become available for market participants. Numerous potential arbitrage opportunities entail simultaneously held long or short positions in puts and calls as well as the underlying futures. Central to the option market is a deterministic relationship between put and call premiums on the same underlying futures contract. This relationship exists because the put, call, and the underlying futures contracts form a portfolio wherein any two of the three instruments can be combined and financed with a riskless bond to replicate the return of the third instrument regardless of the investor's preferences. This relationship is called put-call parity. If the actual premium deviates from the parity price, arbitrage (riskless) profits can be earned.

The original model of put-call parity on stock options was developed by Stoll and later extended by Merton. The model was tested empirically by Klemkosky and Resnick. Option pricing models can be specified for either puts or calls individually. Alternatively, the value of

one type of option (e.g., put) can be deduced from the option pricing model for the other type of option (e.g., call) through put-call parity. The put-call parity relationship binds these markets together.

The most common pricing model for options on commodities is the closed form model developed by Black. This model relates the fair market value of the option premium to five inputs: interest rate, time-to-maturity of the option, underlying futures and exercise price, and standard deviation of the futures price. The standard deviation, which is implied in an option premium, can be obtained by equating the observed option premium and the other four model inputs using an iterative solution technique.¹ If the market is efficient (Fama), the option value should reflect all available information regarding the option and its underlying security. Therefore, participants in efficient markets are not able to trade options or a portfolio of the option with its underlying security to earn riskless profits. If the option model is correct and the market is efficient, the implied standard deviations (*ISD*) derived from puts and calls should be equal and should reflect expectations of the volatility of futures returns over the ensuing time period to maturity.

William W. Wilson is an associate professor, Department of Agricultural Economics, North Dakota State University; H.-G. Fung is an assistant professor, Department of Economics and Finance, University of Baltimore.

Published with the approval of the director of the North Dakota Agricultural Experiment Station, Art. No. 1806.

The authors would like to thank two anonymous reviewers for their helpful comments and suggestions.

¹ Plato (1983, 1985) and others have demonstrated that the interest rate forecast has little impact on the premium. Thus, we treat the observed treasury bill rate as a given parameter in this study similar to other studies (Wilson, Fung, and Ricks).

The purpose of this study is to examine arbitrage possibilities between put and call options on grain futures using two different, yet complementary, tests. In the first test, market efficiency is evaluated using the put-call parity criterion over the duration of trading. If the options markets are efficient and if the model used is appropriate, the *ISD* for puts and calls should be equal. If they are not equal, either the option model used may be inappropriate or market efficiency may be questioned. Although many studies use *ISDs* derived from option pricing models, the appropriateness of the pricing model cannot be completely resolved using this method. In the second part of the paper we examine market efficiency using a test of violation of arbitrage bounds (i.e., put-call parity). If violations of these conditions are observed, riskless profits or market inefficiency can be inferred.

An Ex Ante Volatility Test of Market Efficiency

Put-call parity is evaluated in this section over the duration of trading. *ISDs* are calculated for each strike price and combined using a weighting function. Alternative weighting schemes are presented first, followed by presentation of the data used and empirical results.

Weighting Schemes and the Option Pricing Model

The option pricing model used in this study is based on Black. The call premium, c , can be expressed as

$$(1) \quad c = [FN(d_1) - XN(d_2)]e^{-r(t-t^*)},$$

where

$$d_1 = \frac{\ln(FX) + (\sigma^2/2)(t - t^*)}{\sigma\sqrt{(t - t^*)}}, \text{ and}$$

$$d_2 = d_1 - \sigma\sqrt{(t - t^*)};$$

F is the underlying futures price, X is the strike (exercise) price, r is the riskless interest rate, $(t - t^*)$ is the time-to-maturity, σ^2 is the variance of the futures returns, and $N(\cdot)$ is the cumulative normal density function. The value of the put option is linked to that of the call option through put-call parity (Asay and Wolf). This relation can be expressed as

$$(2) \quad c - p = (F - X)e^{-r(t - t^*)}.$$

From the above equation, the put premium (p) is simplified as

$$(3) \quad p = [XN(-d_2) - FN(-d_1)]e^{-r(t-t^*)}.$$

Examination of the call premium [equation (1)] and put premium [equation (3)], demonstrates that both pricing equations require the same inputs.

Given the four model inputs and the market premium for each (call or put) from equations (1) and (3), the *ISD* in each trading instrument can be estimated. These *ISDs* represent estimates of the *ex ante* volatility of the futures price. Consequently, a separate estimator for *ISD* can be derived for each strike price for each put and call on every trading day. When the *ISDs* from individual strike prices are computed using daily closing option and futures prices, a nonsynchronous measurement bias may occur in the point estimates used to test the null hypothesis.

To minimize this bias, Cox and Rubenstein suggest using a weighted *ISD*, which combines *ISDs* from different strike prices. A number of different weighting schemes are utilized in this study to confirm the robustness of the conclusions. For each day (t) a weighted average of the *ISD* is derived from a portfolio comprised of call options (ISD_j^c) with different strike prices. Similarly, for each day a weighted average of the *ISDs* is derived from a portfolio comprised of put options (ISD_j^p) with different strike prices. The purpose is to test for put-call parity over the duration of the study period by examining whether the weighted *ISD* for calls ($WISD_t^c$) is equal to the weighted *ISD* for puts ($WISD_t^p$).

Given the option premiums and other model inputs, *ISDs* can be computed from equations (1) and (3) for each strike price for both puts and calls using an iterative procedure. Thus, a weighted *ISD* for the option traded at time t can be calculated as follows:

$$(4) \quad WISD_t^c = \sum_{j=1}^N W_{jt}^c ISD_{jt}^c,$$

$$(5) \quad WISD_t^p = \sum_{j=1}^M W_{jt}^p ISD_{jt}^p,$$

where $WISD_t^c$ ($WISD_t^p$) is the weighted *ISD* for call (put) at time t ; $N(M)$ is the number of strikes traded for calls (puts) at time t ; ISD_j^c (ISD_j^p) is the implied standard deviation for a call (put) derived for strike price j ; and W_{jt} is the weighting function used to combine the *ISDs* derived from the different strike prices.

The simplest weighting function applies equal weights to the *ISD* derived from each strike price for individual puts and calls. The equally weighted *ISD* can be expressed as $W_{jt}^c = 1/N$ and $W_{jt}^p = 1/M$. Schmalensee and Trippi argue that results using equal weights are similar to other methods. A number of studies (e.g., Ball and Torous, Beckers, Ramaswamy and Sundaresan, Shastri and Tandon) have concluded that the estimate of *ex ante* volatility derived from the at-the-money option (ATM) are appropriate. In this scheme, the *ISD* for the strike price which is closest to the money is assumed to represent the *ex ante* volatility. This scheme has its own appeal because most trading takes place in at-the-money options. However, some useful information may be neglected in this weighting scheme because options other than those closest to the money are ignored.

Chiras and Manaster suggest using the following weights to combine *ISDs* derived from different strike prices:²

$$(6) \quad W_{jt}^c = (\partial c / \partial ISD_j) \left(\frac{ISD_j}{c} \right) / \left[\sum_{j=1}^N (\partial c / \partial ISD_j) \left(\frac{ISD_j}{c} \right) \right],$$

$$(7) \quad W_{jt}^p = (\partial p / \partial ISD_j) \left(\frac{ISD_j}{p} \right) / \left[\sum_{j=1}^M (\partial p / \partial ISD_j) \left(\frac{ISD_j}{p} \right) \right];$$

where

$$\frac{\partial c}{\partial ISD} = \frac{FT}{2\pi} \exp[-(rT + d_1^2/2)], \text{ and}$$

$$\frac{\partial p}{\partial ISD} = \frac{FT}{2\pi} \exp[-rT + d_1^2/2],$$

where T is $(t - t^*)$, time to maturity.

The logic of the above weighting function is that the option with the largest price elasticity yields the most accurate estimate for futures price volatility; therefore, it is assigned the greatest weight. Other options are included in the average, but the weights are proportional to elasticity. Chiras suggested that weights in proportion to the price sensitivity yield the best estimate

for futures price variation. Thus, the Chiras weighting function is

$$(8) \quad W_{jt}^c = (\partial c / \partial ISD_j) \left[\sum_{j=1}^N \partial c / \partial ISD_j \right],$$

$$(9) \quad W_{jt}^p = (\partial p / \partial ISD_j) \left[\sum_{j=1}^M \partial p / \partial ISD_j \right],$$

where $\partial c / \partial ISD_j$ and $\partial p / \partial ISD_j$ are as defined above.

The Chiras and Manaster method differs from the Chiras method in that the former incorporates the elasticity of future volatility, while the latter only adjusts for the price sensitivity of the volatility. Both of these methods are more complex than the equal or ATM schemes derived above.

Beckers suggests a weighting scheme obtained by minimizing a quadratic function for puts and calls, i.e.,

$$(10) \quad f^c(ISD) = \sum_{j=1}^N W_j^c (c_j - BC)^2,$$

$$(11) \quad f^p(ISD) = \sum_{j=1}^M W_j^p (p - BP)^2,$$

where $BC(BP)$ is the Black model call (put) premium using *ISD* as an input, and $c(p)$ is the actual market call (put) premium. The weighting functions corresponding to Beckers' formulation are

$$W_j^c = (\partial c / \partial ISD_j) / \left[\sum_{j=1}^N \partial c / \partial ISD_j \right], \text{ and}$$

$$W_j^p = (\partial p / \partial ISD_j) / \left[\sum_{j=1}^M \partial p / \partial ISD_j \right].$$

In Beckers' method, selection of *ISD* is different from the other weighting schemes. Given an arbitrary futures volatility and other model inputs, Black call (BC) and put (BP) premiums can be computed using equations (1) and (3). Other values can be computed using alternative futures volatility. Using these values, the quadratic function in equations (10) and (11) can be derived. In Beckers' method the *ISD* is selected that minimizes the quadratic function of the difference between market premiums and its theoretic value is selected.

The simplest weighting schemes either apply equal weight to the *ISDs* derived for each option

² Derivation of $\partial c / \partial ISD$ is available from the authors upon request.

or use the *ISD* derived in the ATM option. The other weighting schemes (Chiras, Chiras and Manaster, Beckers) were designed to give more weight to options whose premium should theoretically be more sensitive to volatility. As a result, options that are far into- and out-of-the money are given the smaller weight. The equal weight scheme assumes that each option is equally important in determining the volatility, and the at-the-money scheme assumes that options other than at-the-money have too much noise to be useful and therefore are given no weight.

For each weighting scheme, we define $\epsilon_t = WISD_t^c - WISD_t^p$, where $t = 1, \dots, n$. If ϵ_t are independently identically distributed random variables with a mean and variance, then the null hypothesis that $H_0: \epsilon_t = 0$ can be tested where $\epsilon_t = \sum_{i=1}^n \epsilon_i / n$. The testing procedure can be performed by invoking the central limit theorem, with the required *t*-statistic defined as

$$(12) \quad t = \frac{\epsilon_t}{S_\epsilon / n},$$

where S_ϵ is the sample estimate of standard deviation of ϵ_t and n is the sample size.³ When the null hypothesis is rejected, we can infer that the option market is not efficient. The maintained hypothesis is that the option model used is appropriate. Violation of this parity condition implies the market forecast of volatility for calls differs from that for puts meaning that market participants would have different forecasts of the same underlying futures price volatility.

Data and Scope of Analysis

Options on selected agricultural futures commenced trading in late 1984. By late 1985 options were being traded on most agricultural futures. For this study four commodities traded at three different exchanges were selected. These commodities include corn and soybeans at the Chicago Board of Trade (CBT) and wheat at the Minneapolis Grain Exchange (MGE) and the Kansas City Board of Trade (KCBT). For each commodity one month was chosen that was consistent with the new crop of the commodity. A single contract month was analyzed (rather than aggregating across contract months) because the nature and extent of informational flows and volatility differ depending on contract month.

Use of options in new-crop months requires traders to forecast volatilities throughout the informational stages of the production cycle.

Daily observations were used, and all of the strike prices traded during the time period were included in the analysis. Strike prices on individual days in which volume was nil were not included in the analysis. The data commenced from about twelve months prior to expiration in the case of the CBT contracts, from January of each year in the case of wheat. The time frame for the latter was slightly shortened because of limited trade in earlier months.

Settlement prices for futures and options were taken from the annual reports of the individual exchanges. The risk-free interest rate was taken from U.S. Treasury bill quotations in the *Wall Street Journal*. Treasury bills were chosen to coincide with the period concurrent with expiration of each option contract (i.e., June, August, October, and November), and converted to an annualized rate, r , for use in the Black model as $r = (10,000/TB)^{360/(t-r^*)} - 1$, where TB is the market price of the treasury bill.

For each strike price, the *ISD* was derived for puts and calls using an iterative solution to the Black pricing model with convergence criterion at ± 0.005 . Given the derived *ISDs* for each strike price, alternative weighting schemes were used to compute an estimate of the *ex ante* market volatility on each day for both puts and calls. The weights and summation for different weighting schemes were derived across all strike prices with nonzero volumes of trade on each day.

Derivation of the *ex ante* volatility using Beckers' weighting scheme was more complex. For each strike, the Black price was calculated assuming alternative values of *ISD* ranging from .07 to .90 by .015. From the Black pricing equation, for example, an option premium was derived for each of the iterated values of *ISD*. These f^c and f^p were evaluated across the alternative values of *ISD*, and the one that minimized f^c and f^p was selected as the measure of the *ex ante* volatility.

The data used in the analysis are described in table 1. Several points of interest arise in comparing volume of trade figures across options and strike prices. First, the distribution of trading across strikes differs for puts and calls. For some options, such as 1987 call in corn, the greatest volume is at the middle strike price, and diminishes for farther in-the-money or out-of-the-money strikes. For other options, the distribution of trade strikes is more sporadic. These results illustrate

³ This testing procedure has been used recently by Johnson.

Table 1. Trading Activity in Selected Options: Total Trading Volume

Strike	December Corn										November Soybeans										Wheat									
	1986					1987					1986					1987					July KCBT					1986				
	Call	Put	Call	Put	Strike	Call	Put	Call	Put	Strike	Call	Put	Call	Put	Strike	Call	Put	Call	Put	Strike	Call	Put	Call	Put	Call	Put	Call	Put	Call	Put
	— ^a	1,293	7,959	23,584	4,793	26,023	18,178	24,544	35,487	21,341	309,038	26	450	1,034	21,048	592	7,025	15,088	220	22	268	—	—	—	—	—	—	—	—	—
140	— ^a	610	4,793	23,584	4,793	26,023	18,178	24,544	35,487	21,341	309,038	26	450	1,034	21,048	592	7,025	15,088	220	22	268	—	—	—	—	—	—	—	—	—
150	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
160	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
170	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
180	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
190	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
200	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
210	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
220	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
230	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
240	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
250	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
Total	142,992	93,769	211,030	412,698	132	488	—	—	—	—	205,501	81,377	340,945	109,428	330	1,985	2,131	5,235	4,549	303	378	307	219	—	—	—	—	—	—	—
Number of trading days	244	246	246	246	246	246	246	246	246	246	246	246	246	246	246	246	246	246	246	246	246	246	246	246	246	246	246	246	246	246
Beginning date	11/12/85	12/1/86	10/15/86	10/15/86	10/15/86	10/15/86	10/15/86	10/15/86	10/15/86	10/15/86	10/15/86	10/15/86	10/15/86	10/15/86	10/15/86	10/15/86	10/15/86	10/15/86	10/15/86	10/15/86	10/15/86	10/15/86	10/15/86	10/15/86	10/15/86	10/15/86	10/15/86	10/15/86	10/15/86	10/15/86

^a Dash indicates not included in analysis.

that the larger in-the-money options have lower trading volume than other options.⁴ A second observation is that in corn and soybeans volume is very unbalanced; the volume in one type of option is substantially greater than the other. As an example, trade in 1986 soybeans calls was two and one-half times that of puts. On the other hand, trading volume was more balanced in the wheat contracts.

Results

The average $ISDs$ for the individual strike prices at each market are presented in table 2. The t -statistic is a test of the null hypothesis that ISD_j^p and ISD_j^c are equal for that particular strike price. In some cases the number of observations was insufficient to conduct the test. The results indicate that the $ISDs$ vary substantially across strike prices. For example, in 1986 corn the ISD_j^p with strikes less than 180, exceeded the corresponding ISD_j^c (i.e., $ISD_j^p > ISD_j^c$ for $j < 180$) while ISD_j^p with strikes greater than 180 ISD_j^p was less than the ISD_j^c (i.e., $ISD_j^p < ISD_j^c$ for $j > 180$). These results suggest that calls were overvalued (undervalued) relative to puts for higher (lower) strike prices. This phenomenon is not apparent in any of the other contracts. The results for soybeans indicate that $ISDs$ for the put and call in 1986 and 1987 are significantly different from each other.

Comparison of the individual $ISDs$ for the 1986 wheat contracts (KCBT and MGE) suggests that the ISD_j^c are all greater than the ISD_j^p for all strikes. This result suggests that puts were undervalued relative to calls during this period. However, for KCBT wheat in 1987 the difference suggests that the puts were overvalued relative to calls (i.e., $ISD_j^p > ISD_j^c$). These results suggest that the extent or nature of arbitrage opportunities may differ substantially in different time periods. In addition, the t -statistic suggests that the $ISDs$ for puts and calls are significantly different in numerous cases. Of the fifty-seven strikes in table 2 where comparison can be made, thirty-two have $ISDs$ for puts and calls that are significantly different indicating violation of the

⁴ Though day-to-day comparisons are necessary to draw this conclusion, it is also reflected in the data in table 1. The reason is likely because options do not have a futures-style margining mechanism. Market makers would routinely offer to buy (sell) in-the-money options at prices less (greater) than the intrinsic value. Prices at which these trades occur would be reflected in the settlement price depending on when the trade occurs and the settlement rules of the respective exchanges.

Table 2. ISDS for Individual Puts and Calls

CBT Corn															
Strike	140	150	160	170	180	190	200	210	220	230	240	250			
December 1986															
Put	.315	.281	.273	.249	.241	.227	.215	.215	.207	.211					
Call	^a	.248	.271	.247	.259	.261	.245	.251	.241	.231					
<i>t</i>	^a	-1.78*	-0.97	-1.50	1.39	2.39*	3.08*	3.92*	3.99*	2.87*					
December 1987															
Put	.346	.252	.242	.235	.232	.238	.276	.289	.289	.308					
Call	.001	^a	.212	.229	.232	.240	.257	.272	.290	.321					
<i>t</i>	^a	-3.38*	-4.76*	0.45	3.66*	3.62*	1.56*	0.97	0.56	.395					
											.334				
											^a				
CBT Soybeans															
Strike	450	475	500	525	550	575	600	625	650						
November 1986															
Put	.189	.165	.170	.172	.190	.180	.282	^a							
Call	.180	.166	.172	.191	.206	.229	.239	.250							
<i>t</i>	-2.78*	-0.56	0.94	3.52*	2.65*	2.33*	-0.60								
November 1987															
Put	.174	.178	.184	.209	.224	.248	.273	.303	.334						
Call	.136	.153	.169	.191	.217	.234	.281	.312	.334						
<i>t</i>	-3.16*	-6.12*	-2.75*	0.22	1.89*	2.55*	2.55*	5.67*	^b						
KCBT Wheat															
Strike	220	230	240	250	260	270	280	290	300	310	320				
July 1986															
Put	^a	.226	.219	.215	.223	.222	.187	.197	.100	.213	^a				
Call	^a	^a	.231	.216	.238	.242	.265	.275	.268	.308	^a				
<i>t</i>	^a	^a	-1.32	-1.69*	-1.99*	-0.11	1.02	0.26	2.09*	^a	^a				
July 1987															
Put	.176	.156	.183	.200	.206	.233	.254	.255	.351	^a	^a				
Call	.120	.199	.159	.189	.192	.179	.206	.241	.263	.286	^a				
<i>t</i>	12.00*	-1.42	-1.07	-3.72*	-1.17	-2.57*	-3.40*	-1.53	-3.90*	^a	^a				
MGE Wheat															
Strike	250	260	270	280	290	300	310	320	330						
September 1986															
Put	^a	.244	.247	.202	.195	.106	.229	^a	^a						
Call	^a	.271	.263	.274	.249	.219	.189	^a	^a						
<i>t</i>	^a	^a	-0.17	0.59	-3.82*	1.25	-1.95*	^a	^a						
September 1987															
Put	.174	.155	.182	.205	.199	.203	^a	^a	^a						
Call	^a	.197	.186	.220	.238	.268	.277	.260	.219						
<i>t</i>	^a	^b	0.10	0.14	0.23	^b	^a	^a	^a						

Note: *t* is test statistic (*t*) for difference between means; asterisk indicates significance at the 10% level.
^a *N* = 0 (no. of observations) in the comparison of differences.
^b *N* = 1 in the comparison of differences.

put-call parity condition. Of the significant comparisons, the ISD^p exceeded the ISD^c in eighteen cases, although the extent of the difference varied across contracts.

The robustness of conclusions on market efficiency is limited when comparing $ISDs$ for individual strike prices. Significant differences may be found between $ISDs$ with some strikes but not with others. Consequently, alternative weighting functions were used for deriving weighted $ISDs$ for puts and calls derived from individual strikes. In addition, the $ISDs$ for the strike price closest to the money (ATM) should incorporate more information than other options.

$ISDs$ based on these different weighting functions are shown in table 3 for each of the contracts. The null hypothesis is that the *ex ante* put and call volatilities are equal. Results of differ-

ent weighting schemes indicate that there appears to be a greater incidence of violation of the put-call parity condition for 1987 than 1986. The frequency of null hypothesis rejection is similar for corn and soybeans in 1986 and 1987. When different contracts are compared, the evidence does not suggest that any one commodity has a greater frequency of efficiency than the others. The frequency of violations in put-call parity is comparable across contracts. In general, the null hypothesis of put-call parity is rejected. With the exception of KCBT wheat, the results indicate there is minimal difference between $ISDs$ derived from puts and calls using the ATM weighting function. This result is consistent with the hypothesis that ATM options reflect more information (e.g., Beckers, Shastri and Tandon) than other studies.

Comparisons across weighting functions are

Table 3. ISD 's Derived Using Alternative Weighting Schemes

	Alternative Weights				
	ATM	Equal	Chiras & Manaster	Chiras	Beckers
CBT Corn					
December 1986					
Put	.226	.228	.227	.223	.144
Call	.228	.217	.239	.224	.439
<i>t</i>	-2.27*	3.94*	6.68*	1.38	16.32*
December 1987					
Put	.232	.237	.237	.241	.154
Call	.233	.239	.250	.238	.584
<i>t</i>	-1.03	1.65*	7.73*	-2.33*	35.25*
CBT Soybeans					
November 1986					
Put	.171	.171	.173	.173	.100
Call	.173	.206	.223	.193	.369
<i>t</i>	1.25	17.37*	26.03*	22.13*	23.08*
November 1987					
Put	.180	.183	.179	.183	.133
Call	.184	.195	.215	.193	.503
<i>t</i>	0.88	8.86*	19.99*	10.88*	27.76*
KCBT Wheat					
July 1986					
Put	.228	.211	.218	.216	.327
Call	.214	.216	.219	.214	.136
<i>t</i>	-2.91*	-0.30	-1.00	-1.55	8.30*
July 1987					
Put	.192	.193	.196	.194	.132
Call	.181	.178	.181	.179	.320
<i>t</i>	-5.73*	-3.69*	-5.02*	-4.89*	8.19*
MGE Wheat					
September 1986					
Put	.220	.212	.217	.216	.182
Call	.213	.233	.233	.233	.214
<i>t</i>	-0.69	1.91*	1.59	1.81*	1.99*
September 1987					
Put	.172	.182	.183	.174	.175
Call	.187	.224	.224	.224	.225
<i>t</i>	-0.18	3.58*	3.63*	4.53*	1.74*

Note: *t* is test statistic (*t*) for difference between means; asterisk indicates significance at the 10% level.

shown in table 4. These comparisons indicate whether different weighting schemes yield different *ISDs*. Despite the criticism that equal weighting schemes give arbitrary weight, the equal weighting function results in *ex ante* estimates of volatility that are not significantly different from the others. In fact, Beckers' is the only weighting scheme significantly different from the others. In the 1986 soybean contract, for example, the *ISD^P* is about .17 for each of the weighting schemes, whereas that for Beckers' is .10; the *ISD^C* ranges from .17 to .22 but is .37 using Beckers'. Similar differences exist in the other contracts except, perhaps, MGE wheat. For calls, the *ISD^C* using ATM was less than *ISD^C* using the other weighting functions except in 1986 corn. In the wheat market, the *ISD^P* using ATM exceeded the *ISD^C* of the other schemes.

Arbitrage Bound Test of Market Efficiency

Market efficiency is evaluated in this section using a test of violation of arbitrage bounds, i.e., put-call parity. Arbitrage bounds are developed first followed by a presentation of results.

Arbitrage Conditions

In the previous section, a number of different weighting schemes were examined to test market efficiency using the put-call parity condition

over the entire duration of trading of a particular expiration month. These *ex ante* tests represent a test of necessary conditions for market efficiency on options. In this section, a test of arbitrage bounds is conducted for individual trading days as opposed to the duration of trading. If the options violate the arbitrage bounds, market inefficiency can be claimed. Transaction costs are also incorporated in the analysis as a more robust test of market efficiency.

Several earlier studies (Ball and Torous, Bodurtha and Courtadon, Ramaswamy and Sundaresan, Ogden and Tucker) have evaluated arbitrage conditions as tests of market efficiency. In this study, similar procedures are adopted in testing the options markets for grains. Because arbitrage conditions are applicable to American options, issues associated with the choice of option pricing model (i.e., using the European model to price American options) are irrelevant when boundary conditions are used as a test of put-call parity and efficiency. Further, the constant interest rate assumption used in the Black model can also be relaxed.

The arbitrage conditions for options on futures contracts have been derived in previous studies (e.g., Ramaswamy and Sundaresan, Ball and Torous). The assumptions made are the following: (a) investors are price-takers in frictionless markets; (b) they prefer more wealth to less; and (c) a market exists for pure discount bonds of every maturity. No assumptions on the stochastic process of options or their underlying as-

Table 4. Tests of Differences across Weighting Schemes

		Equal	ATM	Chiras	Beckers
CBT Corn					
December 1986	Puts	3.64*	4.31*	2.77*	11.57*
	Calls	7.45*	10.39*	8.20*	13.85*
December 1987	Puts	0.39	3.02*	2.05*	10.67*
	Calls	9.25*	10.93*	10.94*	27.37*
CBT Soybeans					
November 1986	Puts	1.45	2.43*	0.36	18.84*
	Calls	18.32*	26.18*	13.52*	14.11*
November 1987	Puts	2.39*	0.10	4.81*	5.97*
	Calls	12.32*	17.62*	15.04*	21.36*
KCBT Wheat					
July 1986	Puts	3.01*	0.20	1.65	7.24*
	Calls	2.74*	3.97*	3.07*	5.94*
July 1987	Puts	1.58	1.47	1.13	5.88*
	Calls	1.52	3.69*	1.83	8.64*
MGE Wheat					
September 1986	Puts	2.11*	2.23*	2.20*	2.87*
	Calls	1.39	1.70	1.48	1.78*
September 1987	Puts	0.17	1.26	2.32*	1.07
	Calls	0.48	0.75	0.04	0.15

Note: Test is for significance in the difference between the weighting scheme shown and that of Chiras and Manaster. The *t*-ratio between the Chiras and Manaster and the respective weighting scheme is shown, and * indicates significance at the 5% level. Signs not shown.

sets (futures or forward contracts) are utilized. Most of the arbitrage conditions are similar to those for options on common stocks (Merton). The arbitrage conditions are presented below.

Arbitrage conditions for call options. The lower bound for the American call is

$$(13) \quad C(t, \tau) \geq \max[O, F(t, t^*) - X],$$

and the upper bound for American call is

$$(14) \quad C(t, \tau) \leq F(t, t^*),$$

where F is the underlying futures price at time t , maturing at t^* ; and C is the American option expiring at time t^* , and $t < \tau \leq t^*$.

Arbitrage conditions for put options. The upper-bound condition for the American put is

$$(15) \quad X \geq P(t, \tau),$$

and the lower bound for American put is

$$(16) \quad P(t, \tau) \geq \max[0, X - F(t, t^*)],$$

where P is the American put price. These arbitrage bounds are valid for (a) any future price distribution and (b) nonstochastic as well as stochastic interest rates.

Parity conditions. The upper parity condition for the put-call relation is

$$(17) \quad C(t, \tau) \leq P(t, \tau) + F(t, t^*) - Xe^{-r\tau},$$

where r is the return on the riskless bond maturing at time t^* . This condition is valid regardless of the future price dynamics and interest rate assumption. The lower parity condition can be expressed as

$$(18) \quad C(t, \tau^*) \geq P(t, \tau) + F(t, t^*)e^{-r\tau} - X.$$

This condition requires a nonstochastic interest rate.

Empirical Analysis and Results

The analysis of arbitrage conditions was conducted on the same contracts and strike prices described above; however, evaluation of some of the conditions required matching data on each particular trading day. Throughout the time series, the volume of trade in either puts or calls, or both, was zero on a number of days. The

evaluation could not be conducted for those days. Thus, the number of trading days in which the violations could be evaluated varied across strike prices and conditions.

In order to profit from a violation of one of the bounds, a trader would take simultaneous positions and would incur transaction costs. The transaction costs are the same as used by Ball and Torous and by Ogden and Tucker on a per contract basis. Adjusting the contract transaction cost by contract size resulted in a round-trip cost of 0.0348¢ per bushel, which was appropriately added or subtracted from the left-hand side of the boundary condition. As an example, to profit from violation of the lower call boundary condition would require taking a long position in the underpriced call, exercising it immediately, and selling the acquired futures. In this transaction, a round trip would be incurred in futures and a one-way trade in calls for a total of .0348¢ per bushel. To profit from violation of one of the put-call parity conditions would require simultaneous positions in (a) writing the overvalued put, (b) buying the undervalued call, (c) borrowing $Fe^{-r\tau}$, and (d) lending X at the interest rate r for time τ .⁵

In each case the percentage of trading days in which violations were incurred and the number of active trading days over which the conditions were calculated. Results for corn are shown in table 5.⁶ As an example, the 200 strike price for 1986 corn had 193 days in which condition 3 could be evaluated. Of those, there were violations of the condition on 9% of the trading days.

As expected there were no violations of the upper call boundary (condition 2). However, the lower call boundary (condition 1) was violated to some extent in nearly all contracts. The largest incidence was in the December 1987 KCBT wheat market in which violations occurred in 7% of the cases. All of the violations of this condition were incurred at lower strike prices. Violations of this condition suggest that arbitrage profits could have been earned by buying the undervalued calls, exercising immediately, and selling the futures.

The upper put boundary was never violated, as expected. However, the lower put boundary was violated frequently. Violation of this con-

⁵ To evaluate the sensitivity of the results to the transaction costs, the boundary conditions were also tested using alternative values ranging downwards to 0¢/bu. The results were very similar to those reported here.

⁶ Results of violations of arbitrage pricing conditions for soybeans and wheat are summarized in this paper and are available from the authors upon request.

Table 5. Violations of Arbitrage Pricing Conditions on CBT Corn

	Strike Price											Total
	140	150	160	170	180	190	200	210	220	230	240	250
December 86												
Condition												
1: $C(t, \tau) \geq \text{Max}[0, F(t, \tau^*) - X]$	50	2	3	2	0	0	0	0	0	0	0	1
2: $C(t, \tau) \leq F(t, \tau^*)$	0	0	0	0	0	0	0	0	0	0	0	0
3: $P(t, \tau) \geq \text{Max}[0, X - F(t, \tau^*)]$	0	0	0	0	0	0	0	0	0	0	0	0
4: $P(t, \tau) \leq X$	0	0	0	0	0	0	0	0	0	0	0	0
5: $C(t, \tau) \leq P(t, \tau) + F(t, \tau^*) - Xe^{-r\tau}$	0	0	0	0	0	0	0	0	0	0	0	0
6: $C(t, \tau) \geq P(t, \tau) + F(t, \tau^*)e^{-r\tau} - X$	0	0	0	0	0	0	0	0	0	0	0	0
December 87												
Condition												
1: $C(t, \tau) \geq \text{Max}[0, F(t, \tau^*) - X]$	50	2	3	2	0	0	0	0	0	0	0	1
2: $C(t, \tau) \leq F(t, \tau^*)$	0	0	0	0	0	0	0	0	0	0	0	0
3: $P(t, \tau) \geq \text{Max}[0, X - F(t, \tau^*)]$	0	0	0	0	0	0	0	0	0	0	0	0
4: $P(t, \tau) \leq X$	0	0	0	0	0	0	0	0	0	0	0	0
5: $C(t, \tau) \leq P(t, \tau) + F(t, \tau^*) - Xe^{-r\tau}$	0	0	0	0	0	0	0	0	0	0	0	0
6: $C(t, \tau) \geq P(t, \tau) + F(t, \tau^*)e^{-r\tau} - X$	0	0	0	0	0	0	0	0	0	0	0	0

Note: Percent of violation is shown and number of active trading days in which the condition could be evaluated is shown in (). * First number for each strike is a percent of violations including transactions cost (7% = .0348/bu.).
a Insufficient data to test condition.

dition suggests that arbitrage profits could be earned by buying the put, exercising it immediately, and buying the futures. The frequency of violations in this condition was greater for the higher-valued strike prices.

Violations of the put-call parity conditions were relatively fewer in corn and soybean contracts. The 1985 MGE wheat contract was violated 16% of the time, but the condition could be evaluated over relatively few days. For the KCBT wheat contract, the violations were as frequent as 38%—60% depending on the strike price. These results in general confirm those presented in the previous analysis and suggest that arbitrage profits could be earned in some contracts and strike prices using a trading rule based on the put-call parity condition. Violation of the upper boundary put-call parity condition would require buying the undervalued put and selling the overvalued call. Opposite positions would be taken if the lower put-call parity bound was violated.

As a generalization, percents of violations aggregated across all conditions excluding the upper bounds are as follows:

Corn	1986	3
	1987	1
Soybeans	1986	4
	1987	1
KCBT wheat	1986	13
	1987	8
MGE wheat	1986	8
	1987	2

These figures illustrate the apparent greater efficiency of the corn and soybean markets relative to that of wheat in terms of violations of the arbitrage boundary conditions. Furthermore, arbitrage opportunities diminished in 1987 as compared to 1986.

The results of arbitrage bound analyses confirm those presented in the previous section which demonstrated that in most cases the *ISDs* for put and calls were significantly different. The frequency of violations of arbitrage bounds in the grain options markets tend to exceed those of other contracts. Ogden and Tucker, for example, showed that British pound currency options had the most frequent violations at 3.1%, while others were violated less than 1% of the time. Ball and Torous observed violations in arbitrage conditions in gold and sugar in 7.56% and 6.03% of the observations, respectively.

Concluding Comments

Two different yet complementary tests on the put-call parity condition for the selected grain

option markets were conducted in this study. The first was a test of the equality of the *ISD* between put and call options over the duration of the study. Results indicated that various measures of *ex ante* volatility of futures price (*ISD*) over the duration of trading differed significantly across put and call options. The second test was to determine the extent to which arbitrage bounds were violated on individual trading days. The analysis demonstrates that riskless profits could likely be earned by traders even after transaction costs.

The underlying motivation of each of these tests is the fundamental put-call parity relationship; however, the formulation of the empirical tests differs. The test of the equality of *ISDs* derived from puts and calls is a test of market efficiency over the entire duration of trading options for a particular expiration month. The test for violations in the arbitrage condition is a direct test of the put-call parity condition (as well as other arbitrage bounds) for individual trading days, as opposed to the duration of trading. In addition, the results of this test are independent of the choice of the option pricing model, and transaction costs are incorporated. Rejection of the null hypothesis using the first test, *vis-à-vis* the arbitrage bound test, is stronger. This finding may be due to bias in the European option model for pricing American options. Because the arbitrage bound test is free from such model bias, it is the preferred procedure for exploring the parity conditions.

Results of these two tests suggest substantial evidence to reject the put-call parity condition in these markets. These findings imply the wheat option market is not as efficient as other option markets (e.g., currency options) and arbitrage profits are likely to exist. This inefficiency may be temporary because of the newness of the grain option market, as is suggested by the reduction in the percentage of violation of arbitrage bounds in 1987 versus 1986 contracts. These findings also could be biased as a result of a nonsynchronization of data. Options and futures settlement prices were used to analyze efficiency or arbitrage opportunities of these markets. Therefore, the reported bias may not be realizable because (a) the bid-ask spread, an important component of transaction cost; (b) the market impact on options and futures prices; and (c) the time lag in execution of a buy and sell order for arbitrage opportunities are all ignored. The above issues can only be resolved via time-stamped intraday transaction data, which are not easily available. Future research using transaction data

to test efficiency in these markets should be encouraged and warranted.

[Received August 1989; final revision received March 1990.]

References

- Asay, M. R. "A Note on the Design of Commodity Option Contracts." *J. Futures Mkts.* 2(1982):1-7.
- Ball, C., and W. Torous. "Futures Options and the Volatility of Futures Prices." *J. Finan.* 41(1986):857-70.
- Beckers, S. "Standard Deviations Implied in Option Prices as Predictors of Future Stock Price Variability." *J. Banking and Finan.* 5(1981):363-81.
- Black, F. "The Pricing of Commodity Contracts." *J. Finan. Econ.* 5(1976):167-79.
- Bodurthier, J., and G. Courtadon. "Efficiency Tests of the Foreign Currency Options Market." *J. Finan.* 41(1986):151-62.
- Chiras, D. "Predictive Characteristics of Standard Deviation of Stock Returns Deferred from Option Prices." Ph.D. thesis, University of Florida, 1977.
- Chiras, D., and S. Manaster. "The Information Content of Options Prices and a Test of Market Efficiency." *J. Finan. Econ.* 7(1978):213-34.
- Cox, J., and M. Rubinstein. *Option Markets*. Englewood Cliffs NJ: Prentice-Hall, 1985.
- Fama, E. *Foundations of Finance*. New York: Basic Books, 1976.
- Johnson, L. "Foreign-Currency Options, Ex-ante Exchange Rate Volatility and Market Efficiency: An Empirical Test." *Finan. Rev.* 21(1986):433-56.
- Klemkosky, R. C., and B. G. Resnick. "Put-Call Parity and Market Efficiency." *J. Finan.* 34(1979):1141-55.
- Merton, R. C. "The Relationship Between Put and Call: Comment." *J. Finan.* 28(1973):183-84.
- Ogden, J., and A. Tucker. "Empirical Tests of the Efficiency of the Currency Futures Options Market." *J. Futures Mkts.* 7(1987):695-703.
- Plato, G. "Theory of Rational Option Pricing." *Bell J. Econ. and Manage. Sci.* 14(1983):141-83.
- . "Valuing American Options on Commodity Futures Contracts." *Agr. Econ. Res.* 37(1985):1-14.
- Ramaswamy, K., and S. M. Sundaresan. "The Valuation of Option on Futures Contracts." *J. Finan.* 40(1985):1319-40.
- Schamlensee, R., and R. Trippi. "Common Stock Volatility Expectation Implied by Option Premia." *J. Finan.* 33(1978):129-47.
- Shastri, K., and K. Tandon. "On the Use of European Models to Price American Options on Foreign Currency." *J. Futures Mkts.* 6(1986):93-101.
- Stoll, H. R. "The Relationship Between Put and Call Option Prices." *J. Finan.* 24(1969):802-24.
- Wilson, W., H. G. Fung, and M. Ricks. "Option Price Behavior in Grain Futures Markets." *J. Futures Mkts.* 8(1988):47-65.
- Wolf, A. "Fundamentals of Commodity Options on Futures." *J. Futures Mkts.* 2(1982):391-408.

Production, Hedging, and Speculative Decisions with Options and Futures Markets

Harvey Lapan, Giancarlo Moschini, and Steven D. Hanson

This paper analyzes production, hedging, and speculative decisions when both futures and options can be used in an expected utility model of price and basis uncertainty. When futures and option prices are unbiased, optimal hedging requires only futures (options are redundant). Options are used together with futures as speculative tools when market prices are perceived as biased. Straddles are used to speculate on beliefs about price volatility and to hedge the futures position used to speculate on beliefs about the expected value of the futures price. Mean-variance analysis in general is not consistent with expected utility when options are allowed.

Key words: futures markets, hedging, options, price uncertainty, risk.

One extension of Sandmo's expected utility model of the competitive firm under price uncertainty considers the use of futures or forward contracts. Danthine; Holthausen; and Feder, Just, and Schmitz show that without basis uncertainty the optimal output level is not affected by price risk; also, with an unbiased futures price, the optimal hedging level of the competitive firm is the full hedge, while a biased futures price will result in a partly speculative hedge. Related works include Batlin, who allows for basis risk in the form of imperfect time hedging; Paroush and Wolf, and Antonovitz and Nelson, who consider basis risk with the simultaneous availability of futures and forward contracts; Grant, Honda, Losq, Newbery and Stiglitz, and Rolfo, who allow for production uncertainty; Chavas and Pope, who allow for production uncertainty and hedging costs; and Karp, who considers the problem in a dynamic setting.

This paper provides a further extension of this analysis by allowing options as a means of coping with price risk. With the introduction of commodity options on futures for many commodities in the 1980s, this problem appears rel-

evant to a number of production settings, especially in agriculture. Specifically, this paper considers the simultaneous choice of a production level and of hedging levels of futures and options within the general expected utility model. The model allows for basis uncertainty, but the production process is assumed to be nonstochastic.

Optimal hedging when options are available is considered by Wolf in a linear mean-variance framework without including production decisions.¹ The mean-variance framework has been employed in a number of risk management studies. Under certain assumptions, this framework is consistent with expected utility maximization (Meyer, Robison and Barry). However, the inclusion of commodity options in a decision maker's portfolio leads to a violation of the two main conditions for a mean-variance representation of expected utility. First, options truncate the probability distribution of price so that the argument of the utility function, profit or wealth, is not normally distributed even if the random price is normal. Second, the use of options generally means that the argument of utility is not monotonic in the random attributes. Thus, relaxing the mean-variance framework appears desirable to analyze options in a hedging problem.

Harvey Lapan and Giancarlo Moschini are a professor and an assistant professor, respectively, Department of Economics, Iowa State University. Steven D. Hanson is an assistant professor, Department of Agricultural Economics, Michigan State University.

This is Journal Paper No. J-14133 of the Iowa Agriculture and Home Economics Experiment Station, Project No. 2953.

The authors thank Robert Myers and the *Journal* reviewers for their constructive comments.

¹ Wolf also reports simulation results based on the logarithmic utility function.

The paper is organized as follows. A model of production and hedging with both futures and put options is formulated. Some general results are derived for the pure hedging case in which producer price expectations agree with those embodied in the market price of futures and options. This is followed by an analysis of how changes in asset prices (or expectations) affect optimal portfolios under constant absolute risk aversion (CARA). Next, the model is reformulated in terms of futures and straddles. This reformulation highlights the impact that individual beliefs have on speculative decisions, particularly on the use of options, and illustrates the limitations of mean-variance analysis. The concluding section summarizes the main contributions of the paper.

Production and Hedging with Futures and Options

The notation is defined as follows: y is the output quantity produced, x is the futures quantity sold, z is the put option quantity sold, p is the futures price at the end of the period, f is the futures price at the beginning of the period, b is the local cash price (including basis risk) at the end of the period, r is the put option price (premium), k is the strike price, v is the terminal value of a put option, π is the profit at the end of the period, and \sim denotes a random variable. Because one can construct a synthetic call using futures and puts, attention is restricted to put options only. Also, for simplicity, only one available strike price for the option is considered.

The random end-of-period profit of the firm using both futures and put options can then be written as²

$$(1) \quad \pi = \tilde{b}y - c(y) + (f - \tilde{p})x + (r - \tilde{v})z,$$

where $c(y)$ is a strictly convex cost function dual to a concave production function,³ and \tilde{v} is the terminal value of a put option, defined as

$$(2.1) \quad \tilde{v} = 0 \quad \text{if } p \geq k$$

$$(2.2) \quad \tilde{v} = k - \tilde{p} \quad \text{if } p < k,$$

where p is a realization of the random variable \tilde{p} .

The producer's utility is a strictly concave function defined over profit, that is $u = u(\pi)$ where π is given by (1). Thus, the individual is risk averse, but no other restrictions are placed on his/her preferences. The problem is to choose (y, x, z) to maximize expected utility; that is,

$$(3) \quad \max_{y,x,z} \mathcal{L} = E[u(\pi)],$$

where E denotes the mathematical expectation operator.

The first-order conditions (FOC) require $\mathcal{L}_y = \mathcal{L}_x = \mathcal{L}_z = 0$, where the subscripts to \mathcal{L} denote arguments of partial differentiation, that is,⁴

$$(4.1) \quad E[\tilde{u}'(\tilde{b} - c')] = 0$$

$$(4.2) \quad E[\tilde{u}'(f - \tilde{p})] = 0$$

$$(4.3) \quad E[\tilde{u}'(r - \tilde{v})] = 0,$$

where $\tilde{u}' = du/d\pi$, and $c' = dc/dy$.

To characterize the solution of these equations it is necessary to be specific about the relationship between local cash price and futures prices. Following Benninga, Eldor, and Zilcha, and others, the cash price is written as a linear function of the futures price:

$$(5) \quad \tilde{b} = \alpha + \beta\tilde{p} + \tilde{\theta},$$

where \tilde{p} and $\tilde{\theta}$ are independently distributed and $E(\tilde{\theta}) = 0$. Because of the definition in (2), equation (5) also uniquely defines the relationship between the terminal value of the put option and the cash price. Using (5), the FOC in (4) can be rewritten as

$$(6.1) \quad E[\tilde{u}'(\alpha + \beta f + \tilde{\theta} - c')] = 0$$

$$(6.2) \quad E[\tilde{u}'(f - \tilde{p})] = 0$$

$$(6.3) \quad E[\tilde{u}'(r - \tilde{v})] = 0,$$

where (6.1) uses (4.2) rewritten as $E[\tilde{u}'(\beta f - \beta\tilde{p})] = 0$.

The Case of Unbiased Prices

For any given utility function, the solution of equations (6) will depend crucially on the decision maker's perception of the futures price

² Input prices and the option premium are implicitly compounded to the end of the period using the (constant) market interest rate, so that all monetary variables in (1) are commensurable.

³ Because choosing a profit-maximizing level of inputs is equivalent to choosing a level of output when production is nonstochastic, in the production/hedging problem y is the decision variable of the producer. The effects of exogenous variables on the optimal input levels could be obtained using the (nonstochastic) conditional input demand functions implied by the cost function via the derivative property.

⁴ Because the utility function is strictly concave and the cost function is strictly convex in output, the second-order conditions are satisfied.

and option value distribution relative to the prices f and r . To account for this, it is convenient to define the notion of price bias.⁵

Specifically, if the producer's expectation of the end-of-period futures price equals the price of a futures contract, i.e., $\bar{p} \equiv E(\bar{p}) = f$, then the futures price is unbiased. Similarly, the expected (gross) return from the option position, $\bar{v} \equiv E(\bar{v})$, is

$$(7) \quad \bar{v} = \int_0^k (k - p) \psi(p) dp,$$

where $\psi(p)$ is the density function of the distribution of price as perceived by the producer. Following Black and Scholes, general equilibrium option pricing formulae assign a price to the option which, in our framework, is equivalent to the expected returns from the option (i.e., risk attitudes do not matter). Thus, the producer will perceive options to be fairly priced if $r = \bar{v}$, and in this case the option price is unbiased.

To analyze the solution, the strategy is first to consider the benchmark case of unbiased prices. The effects of biased prices are investigated in terms of comparative statics results from this benchmark. For unbiased prices ($\bar{p} = f$ and $\bar{v} = r$), and taking the nonrandom elements out of the expectation operator, the first-order conditions (6) can be rewritten as

$$(8.1) \quad \text{Cov}(\bar{u}', \bar{\theta})/E\bar{u}' = c' - \alpha - \beta f,$$

$$(8.2) \quad \text{Cov}(\bar{u}', \bar{p}) = 0,$$

$$(8.3) \quad \text{Cov}(\bar{u}', \bar{v}) = 0,$$

where $\text{Cov}(\cdot, \cdot)$ denotes the covariance operator.

Consider first the solution for the optimal futures and option positions for any given level of output, which is obtained by solving (8.2) and (8.3) for x and z conditional on y . Because of the dependence in (5) the random profit can be written as

$$(9) \quad \bar{\pi} = \pi_0 + \bar{\theta}y + \bar{p}[\beta y - x] - \bar{v}z,$$

where $\pi_0 = [\alpha y - c(y) + fx + rz]$ is the nonstochastic component of profit. Now consider $x = \beta y$ and $z = 0$ as a candidate solution. With these levels of hedging instruments, the only

randomness left in $\bar{\pi}$ and in \bar{u}' is due to $\bar{\theta}$. Because $\bar{\theta}$ and \bar{p} are by assumption independently distributed, then $\text{Cov}(\bar{u}', \bar{p}) = 0$ and $\text{Cov}(\bar{u}', \bar{v}) = 0$ so that $x^* = \beta y$ and $z^* = 0$ solve equations (8.2) and (8.3). Because the second-order conditions hold globally, this solution solves the expected utility-maximization problem.

Given this optimal choice of hedging instruments, equation (8.1) will solve for the optimal level of output y^* . Because in this case random profit reduces to $\bar{\pi} = (\alpha + \beta f + \bar{\theta})y - c(y)$, the choice of output level reduces to the standard problem of the competitive firm under output price uncertainty, where the random price is $(\alpha + \beta f + \bar{\theta})$. Using known results of this model, under risk aversion, production takes place at a point at which marginal cost is lower than the expected price (with optimal hedging), i.e., $c'(y^*) < \alpha + \beta f$, indicating that a portion of price risk resulting from the basis cannot be hedged away.

Because there is some residual uncertainty concerning the local cash price, the degree of risk aversion also influences optimal output. Specifically, the output level y^* is inversely related to the degree of risk aversion (Baron). Finally, a *ceteris paribus* increase in nondiversifiable basis uncertainty (a mean-preserving spread of $\bar{\theta}$) will in general decrease the optimal output level, a sufficient condition being that the Pratt measure of absolute risk aversion is decreasing in profit (Ishii). These conclusions are summarized in the following:

Result 1. When futures and options prices are unbiased and cash and futures prices are related as in (5), then (a) a fraction β of production is hedged in the futures market, (b) options are not used as hedging instruments, (c) and the portion of nondiversifiable basis risk affects the production level.

The absence of options from the optimal hedge may seem counterintuitive and warrants clarification. In the absence of futures and options positions, the risk faced by the producer depends upon the distribution of cash prices which, from (5), is assumed to be a linear function of the end-of-period futures price \bar{p} plus an orthogonal component $\bar{\theta}$. Neither the futures contract nor the put option can provide any hedge against the basis risk that is independent of \bar{p} . However, because the remaining hedgeable risk is linear in \bar{p} , it follows that a futures contract (which yields a pay-off that is also linear in \bar{p}) must provide a better hedge than an option contract (the pay-off of which is nonlinear in \bar{p}). Thus, in this context the option has no value as

⁵ This definition of bias illustrates the producer's beliefs about the price distribution and does not warrant any implication about market efficiency.

a hedging instrument if futures contracts are also present.⁶

The optimal hedge ratio for futures derived above, $x^*/y = \beta$, is the same as that derived under similar conditions for the case of futures only (Benninga, Eldor, and Zilcha; Kahl). This optimal hedge ratio satisfies the condition $\beta = \text{Cov}(\tilde{b}, \tilde{p})/\text{Var}(\tilde{p})$ and thus could be estimated by a linear regression of cash on futures prices.

Additional results for the general solution may be obtained under special conditions. First, if there is no orthogonal basis risk ($\tilde{\theta} \equiv 0$), then the FOC of equation (8.1) reduces to $c'(y^*) = \alpha + \beta f$. This means that the optimal output level is independent of the distribution of \tilde{p} , i.e., production and hedging/speculative decisions are separated. Second, when the basis ($\tilde{b} - \tilde{p}$) and the futures price \tilde{p} are independent (i.e., $\beta = 1$), the optimal hedge is the full hedge in the futures market, although in this case production and hedging decisions are not necessarily separated.

A separation result can also occur under our assumptions concerning basis risk if the utility function is of the CARA type. Assume a negative exponential utility function $\tilde{u} = -\exp(-A\tilde{\pi})$, where $A \equiv -u''/u'$ is the (constant) Pratt coefficient of absolute risk aversion, such that $\tilde{u}' = A \exp(-A\tilde{\pi})$. Because \tilde{p} and $\tilde{\theta}$ are independently distributed, it is verified that for CARA equation (6.1) can be written as

$$(10) \quad \frac{\int \theta \exp(-A\theta y) h(\theta) d\theta}{\int \exp(-A\theta y) h(\theta) d\theta} = c'(y) - \alpha - \beta f,$$

where $h(\theta)$ is the density function of $\tilde{\theta}$. Thus, the optimal output level y^* that solves equation (10) will not be affected by parameters of the distribution of \tilde{p} . This can be summarized in the following:

Result 2. If either (a) there is no basis uncertainty or (b) there is a CARA utility function and cash and futures prices are related as in (5), then there is separation between production and hedging (speculative) decisions.

This separation result means that, when the producer believes that the current futures price is a biased indicator of the end-of-period futures

price, it is more efficient for the producer to speculate on this disparity through portfolio decisions than through production decisions. This result with no basis uncertainty is essentially the same as those of Danthine; Holthausen; and Feder, Just, and Schimtz. The separation result with linear basis risk is illustrated by (10). The general case of (8.1) shows that, at the optimal output level, the marginal cost is less than the expected return from an optimally hedged position by an amount which reflects the risk premium of the unhedgeable risk [the left-hand-side of (8.1)]. As shown in (10), under CARA this risk premium is independent of that portion of price risk which is orthogonal to $\tilde{\theta}$, leading to separation between production and hedging decisions.

The Role of Expectations

Result 1 above describes the hedging decisions of a producer whose expectations agree with the market forecasts as displayed by the futures price and option premium. When this condition is relaxed, the optimal hedging rule $x^* = \beta y$ and $z^* = 0$ is modified. A convenient way to model the divergence of individual expectations from the market expectations, as aggregated in the futures price f and in the option price r , is to let f and r change while holding the producer's subjective distribution of \tilde{p} unchanged.⁷

Consider first hedging decisions conditional on the output level y . In this case equations (4.2) and (4.3) will solve for x^* and z^* . Totally differentiating these FOC and solving yields

$$(11) \quad \begin{bmatrix} dx \\ dz \end{bmatrix} = \frac{1}{\Delta} \begin{bmatrix} -\mathcal{L}_{zz} & \mathcal{L}_{zx} \\ \mathcal{L}_{xz} & -\mathcal{L}_{xx} \end{bmatrix} \cdot \begin{bmatrix} \mathcal{L}_{zf} & \mathcal{L}_{zr} \\ \mathcal{L}_{xf} & \mathcal{L}_{xr} \end{bmatrix} \begin{bmatrix} df \\ dr \end{bmatrix},$$

where

$$\mathcal{L}_{xx} = E[\tilde{u}''(f - \tilde{p})^2],$$

$$\mathcal{L}_{xz} = \mathcal{L}_{zx} = E[\tilde{u}''(r - \tilde{v})(f - \tilde{p})],$$

⁶ There are circumstances in which both futures and options may be useful hedging devices. In general, these cases will display a distribution of profit which is nonlinear in the random price. Examples of these situations may involve a nonlinear basis relationship or the presence of production risk which is not orthogonal to price risk. The characterization of these cases is left for further research.

⁷ This is consistent with the assumptions that f and r are exogenously given to the producer. Alternatively, one could hold the prices f and r constant and investigate how changes in the subjective distribution affect optimal hedging. Because the assumptions do not rule out distributions with more than two moments, this alternative may be impractical.

$$\mathcal{L}_{zz} = E[\tilde{u}''(r - \bar{v})^2],$$

$$\mathcal{L}_{xf} = E[\tilde{u}''(f - \bar{p})x] + E\tilde{u}',$$

$$\mathcal{L}_{xr} = E[\tilde{u}''(f - \bar{p})z],$$

$$\mathcal{L}_{zf} = E[\tilde{u}''(r - \bar{v})x],$$

$$\mathcal{L}_{zr} = E[\tilde{u}''(r - \bar{v})z] + E\tilde{u}', \text{ and}$$

$$\Delta = \mathcal{L}_{xx}\mathcal{L}_{zz} - (\mathcal{L}_{xz})^2.$$

Under CARA, and using the first-order conditions (4.2) and (4.3), it is verified that $\mathcal{L}_{xf} = \mathcal{L}_{zr} = E\tilde{u}'$ and $\mathcal{L}_{xr} = \mathcal{L}_{zf} = 0$. Under CARA, therefore, equation (11) reduces to

$$(12) \quad \begin{bmatrix} dx \\ dz \end{bmatrix} = \frac{E\tilde{u}'}{\Delta} \begin{bmatrix} -\mathcal{L}_{zz} & \mathcal{L}_{zx} \\ \mathcal{L}_{xz} & -\mathcal{L}_{xx} \end{bmatrix} \begin{bmatrix} df \\ dr \end{bmatrix}.$$

From the second-order conditions we know that $\mathcal{L}_{xx} < 0$, $\mathcal{L}_{zz} < 0$, and $\Delta > 0$. It can also be proved that under CARA $\mathcal{L}_{zx} = \mathcal{L}_{xz} > 0$.⁸ Hence, (12) yields:

Result 3. Under CARA, a *ceteris paribus* change in the futures price leads to a change in futures and options sold in the same direction, i.e., $\partial x^*/\partial f > 0$ and $\partial z^*/\partial f > 0$. A *ceteris paribus* change in the option premium leads to a change in futures and options sold in the same direction, i.e., $\partial x^*/\partial r > 0$ and $\partial z^*/\partial r > 0$.

The relative magnitude of the comparative statics effects also provides some insight on how futures and options are used to exploit information on futures and options bias. First, (12) implies that $\partial x^*/\partial r = \partial z^*/\partial f$ because $\mathcal{L}_{zx} = \mathcal{L}_{xz}$. Also, from (12) we obtain

$$(13) \quad dz - dx = (E\tilde{u}'/\Delta) [(\mathcal{L}_{zz} + \mathcal{L}_{xx})df - (\mathcal{L}_{zx} + \mathcal{L}_{xz})dr].$$

Under CARA $(\mathcal{L}_{zz} + \mathcal{L}_{xx}) > 0$ and $(\mathcal{L}_{zx} + \mathcal{L}_{xz}) < 0$.⁹ Thus we can conclude:

Result 4. Under CARA, a *ceteris paribus* change in futures or options price results in a larger change in the option position than in the futures position, with changes in r implying the largest changes. That is: $\partial z^*/\partial r > \partial x^*/\partial r = \partial z^*/\partial f > \partial x^*/\partial f > 0$.

The comparative statics of results 3 and 4 should be interpreted with care because they consider the effects of only r or f changing. When the producer differs from the market in terms of his/her perception of the dispersion of price but not in the expected value, only the option price

will be perceived as biased, and the comparative statics of a change in r applies. Thus, for example, if one started from the unbiased solution of the basic independence case $x^* = y$ and $z^* = 0$, then an increase in r (i.e., the market is overstating the volatility of \bar{p} from the individual point of view) would lead to a partly speculative futures open position $x^* > y$ such that more futures are sold. At the same time a position $z^* > 0$ in the option market is also open, such that some put options are also sold. Because the change in option position is larger than the change in the futures position by result 4, the resulting pay-off of the speculative position resembles a short straddle, a strategy which is deemed useful for speculating on beliefs about the volatility.¹⁰

Alternatively, changes in the producer's beliefs concerning expected price would bias both futures and option prices. If, at the same time, the producer perceives an offsetting change in the dispersion of price, then only the futures price may be perceived as biased. In this case the comparative statics of a change in f applies. However, this special case may not be very interesting; rather, one may want to know the effects of changing the futures price f when volatility is perceived unchanged. In this case the option premium must be allowed to change when the price of futures changes. A way of doing that would be to let $dr = \delta df$, where the coefficient δ is known as the "delta" of the option (Cox and Rubinstein), and to pursue the comparative statics analysis in terms of total derivatives. An alternative and more fruitful analysis involves reformulating the problem in terms of futures and straddles. This approach is pursued in the next section.

Speculative Hedging with Options

The analysis so far has shown little role for options as a hedging instrument. As long as the exogenous risk is linear in price, the futures market provides a perfect hedge. Moreover, the individual speculative decisions depend upon market prices (f and r) and his/her subjective beliefs concerning the "fair" value of these prices. If these prices are perceived as biased, then a speculative position which includes the use of options may occur.

⁸ Proof of this result is available from the authors upon request.

⁹ Proof of this result is available from the authors upon request.

¹⁰ A short straddle position is obtained by selling one futures and two put options, or equivalently by selling one put and one call options. See Cox and Rubinstein for more details on this and other strategies.

The results concerning speculative decisions are obscured because the futures and put contracts are partial substitutes (a futures equals a long call and a short put). As pointed out before, the bias in the futures price depends on expected price, while the bias in option premiums depends on both expected price and volatility. Thus, changes in price expectations have a direct impact on the expected return (or bias) of both instruments.

The analysis of speculative decisions can be sharpened if the individual speculates using futures contracts and straddles. A long (short) straddle is a combination of a long (short) call and a long (short) put at the same strike price. Because a synthetic call can be constructed using futures and puts, a straddle can also be constructed using futures and puts. Thus, recasting the analysis with this instrument does not entail a change in the choices available to the individual.

The main determinant of the straddle price is the volatility of the end-of-period futures price. From the individual perspective, changes in expectations concerning the futures price affect the bias in the futures but have little effect on the bias in the straddle, while changes in beliefs about the dispersion of price (volatility) affect only the bias in the straddle.

Because basis risk has been investigated earlier, the analysis is simplified by assuming no basis risk ($b \equiv \bar{p}$) and by concentrating on the speculative decisions given a (fully hedged) output level. To this end, let $q \equiv (y - x)$ denote the speculative (open) futures market position ($q > 0$ is long), s be the straddle position ($s > 0$ is a long straddle), t be the market price of a long straddle with strike price k (the premium of the put plus the premium of the call), and \bar{w} be the payoff of the straddle position. The profit defined in terms of open futures and straddle positions then is written as

$$(14) \quad \bar{\pi} = fy - c(y) + (\bar{p} - f)q + (\bar{w} - t)s.$$

The payoff of the straddle is given by the absolute value function,

$$(15) \quad \bar{w} = |\bar{p} - k|.$$

It is convenient to let $\bar{p} \equiv \bar{p} + \bar{e}$, where \bar{e} is a random variable with density function $g(e)$ and satisfying $E(\bar{e}) = 0$. Also, assume that the strike price is chosen equal to the expected price, so that $k = \bar{p}$ and $\bar{w} = |\bar{e}|$. If one maximizes the expected utility of profit as given in (14) conditional on the output level, the first-order conditions are

$$(16.1) \quad E[\bar{u}'(\bar{p} - f + \bar{e})] = 0, \text{ and}$$

$$(16.2) \quad E[\bar{u}'(|\bar{e}| - t)] = 0.$$

For the results that follow, assume that the individual perceives the price to be symmetrically distributed, i.e., $g(e) = g(-e)$. Then,

$$(17) \quad \int_{e < 0} F(e)g(e)de = \int_{e > 0} F(-e)g(e)de$$

for any function $F(e)$. For notational convenience, define¹¹

$$(18) \quad E^+[J(e)] \equiv 2 \int_{e > 0} J(e)g(e)de$$

for any function $J(e)$. Using this symmetry assumption, the FOCs can be expressed in terms of the realizations $e \geq 0$ only. In particular, equations (16) can be rewritten as

$$(19.1) \quad E^+[K(\bar{e})\bar{e}] + (\bar{p} - f)E^+[L(\bar{e})] = 0,$$

$$(19.2) \quad E^+[L(\bar{e})(\bar{e} - t)] = 0,$$

where $K(\bar{e}) \equiv [u'(\pi(\bar{e})) - u'(\pi(-\bar{e}))]$ and $L(\bar{e}) \equiv [u'(\pi(\bar{e})) + u'(\pi(-\bar{e}))]$.¹²

Note that $E^+[L(\bar{e})] > 0$ and, assuming risk aversion ($u'' < 0$), for $e > 0$:

$$(20) \quad K(e) \geq 0 \text{ as } q \leq 0.$$

Thus, if the futures price is unbiased ($\bar{p} = f$) then $q^* = 0$ for (19.1) to hold. If $\bar{p} > f$, then (19.1) requires $E^+[K(\bar{e})\bar{e}] < 0$, which implies $q^* > 0$ in view of (20). Similarly, if $\bar{p} < f$, then $E^+[K(\bar{e})\bar{e}] > 0$, which requires $q^* < 0$. This can be summarized as follows:

Result 5. Under risk aversion and with a symmetric distribution of price, the qualitative optimal speculative futures position depends only upon the bias in the futures price; i.e., $q^* \geq 0$ as $\bar{p} \geq f$. In particular, this speculative futures position is independent of the bias in the straddle price.

The reformulation of the asset mix in terms of futures and straddles allows a clearer representation of how changes in expected price or implied volatility affect optimal speculative de-

¹¹ The normalization factor 2 in (18) is the reciprocal of the probability of positive realizations of \bar{e} (which is 0.5 because of the symmetry assumption), so that E^+ is properly interpreted as an expected value.

¹² $u'(\pi(\bar{e}))$ denotes the marginal utility evaluated at the profit level associated with a realization of the random variable \bar{e} , while $u'(\pi(-\bar{e}))$ denotes the marginal utility evaluated at the profit level associated with the negative of the same realization of the random variable \bar{e} .

cisions. This reformulation also clarifies the comparative statics of results 3 and 4. For example, a perceived decline in the dispersion of \bar{p} means that the producer/speculator views both puts and calls as overpriced and hence wishes to sell both, i.e., sell a straddle. If the problem is formulated in terms of futures and puts, then the sale of a (synthetic) straddle can be achieved by selling one futures and two puts. Thus, the use of futures to speculate on the dispersion of price emerging from result 3 is purely a function of constructing a straddle position.

The first part of this paper showed that options are not useful in hedging exogenous price risk, at least under the assumption of linear price risk. With the reformulation of the model using the straddle, however, we can show that options are desirable instruments to hedge the risk assumed by the agent because of an open speculative position in futures. For example, if the agent has a long futures position, then profit (and utility) will be low when p is low. In this situation, a long straddle that raises income in these states may be desirable. The same argument indicates that a long straddle may be useful when the agent chooses to short the futures contract.

Specifically, the FOC (19.1) determines the sign of q^* , which depends on the perceived bias in the futures price. The sign of s^* then is based on the remaining FOC (19.2). If the straddle is unbiased, then $\bar{w} = t$ where $\bar{w} \equiv E(\bar{w}) = E[\bar{e}] = E^+[\bar{e}]$, the last equality following from the symmetry assumption. Thus, the FOC (19.2) can be rewritten as $E^+\{[L(\bar{e}) - L(\bar{w})](\bar{e} - \bar{w})\} = 0$, where $L(\bar{w})$ denotes the value of $L(e)$ at $e = \bar{w}$. Now, $L(e)$ is an increasing function of e at $s = 0$ if we assume $u''' > 0$ (as implied by non-increasing absolute risk aversion) because $dL/de = [u''(\bar{e}) - u''(-\bar{e})]q$ at $s = 0$, and $d^2L/de^2 = [u'''(\bar{e}) + u'''(-\bar{e})]q^2 > 0$ for $q \neq 0$. Thus, $\text{sgn}[L(\bar{e}) - L(\bar{w})] = \text{sgn}[\bar{e} - \bar{w}]$, so that the integrand in (19.2) is positive for all $e \neq \bar{w}$ and equal to zero at $e = \bar{w}$. Hence, with an unbiased straddle price at $s = 0$ the FOC in (19.2) is positive. Given the second-order conditions, s^* must also be positive for (19.2) to vanish.

For an unbiased straddle price, the long straddle position will not fully offset the open futures position. To see this note that, from (19.2), $E^+[L(\bar{e})(\bar{e} - t)] = \text{Cov}^+[L(\bar{e}), (\bar{e} - t)] = 0$ because $E^-(\bar{e} - t) = 0$. From the definition of $L(e)$, $dL/de = u''(\bar{e})(s + q) + u''(-\bar{e})(s - q)$. Because $L(e)$ cannot be monotonic in e (if this were the case then $\text{Cov}^+[L(\bar{e}), (\bar{e} - t)] \neq 0$), it follows that $(s + q)$ and $(s - q)$ must have opposite sign. Thus, $(s + q)(s - q) = s^2 - q^2 < 0$, which

implies $|s| < |q|$. For example, if $\bar{p} > f$, then $q^* > s^* > 0$, so that the resulting position is not equivalent to simply being a long call.

The analysis above has shown that a long straddle provides insurance against adverse outcomes induced by the open futures position. This conclusion is summarized in the following:

Result 6. Under nonincreasing absolute risk aversion, and with unbiased straddle price and a symmetric distribution of the spot price, the optimal straddle position will always be long when there is an open speculative futures position (either long or short); also, the straddle position will always be smaller than the open futures position, i.e. $0 < s^* < |q^*|$.

Results 5 and 6 together imply that the speculative position induced by a perceived bias in the futures price will have a nonlinear payoff whose shape is intermediate between that generated by only puts or calls and that generated by futures only. This nonlinear payoff will generate a positively skewed profit distribution even if the price distribution is symmetric. Indeed, the condition of nonincreasing risk aversion (or more generally $u''' > 0$) used above is equivalent to assuming a preference for positive skewness (Tsiang).

Note, however, that preference for skewness is not sufficient to justify using only puts or calls to generate a nonlinear payoff. For example, if the agent's expectation of the end-of-period futures price is higher than f but he/she agrees with the implied price volatility of the market, a viable speculative action would seem to involve buying call options. However, if the agent views the call as underpriced, he/she must simultaneously view the put as overpriced. Hence, buying a call and selling a put, i.e., buying futures, is a superior speculative device. Options in a pure sense (straddle) are used only to hedge the risk assumed by this open position.

Comparative statics results concerning the impact of changes in futures and option prices on the optimal portfolio of the individual can be obtained for a CARA utility function. Totally differentiating (16), using the FOC and the constant coefficient of absolute risk aversion $A \equiv -u''/u'$, and solving yields

$$(21) \quad \begin{bmatrix} dq \\ ds \end{bmatrix} = \frac{Eu'}{\Delta} \begin{bmatrix} \mathcal{L}_{ss} & -\mathcal{L}_{qs} \\ -\mathcal{L}_{qs} & \mathcal{L}_{qq} \end{bmatrix} \begin{bmatrix} df \\ dt \end{bmatrix},$$

where \mathcal{L}_{ij} are the elements of the Hessian matrix for the maximization problem, and hence $\mathcal{L}_{qq} < 0$, $\mathcal{L}_{ss} < 0$, and $\Delta = [\mathcal{L}_{qq} \mathcal{L}_{ss} - (\mathcal{L}_{qs})^2] > 0$. Furthermore, with the assumption of symmetry

in $g(e)$, $\mathcal{L}_{qs} \geq 0$ as $q \geq 0$.¹³ Hence: $\partial q^*/\partial f < 0$, $\partial s^*/\partial t < 0$, and $\partial q^*/\partial t = \partial s^*/\partial f \geq 0$ as $q^* \leq 0$. This can be summarized in the following:

Result 7. Under CARA and with symmetrically distributed prices, a *ceteris paribus* decrease in the current futures price will lead to an increase in the net long futures position; the net long straddle position will increase (decrease) if the net futures position is long (short). Similarly, a *ceteris paribus* decrease in the straddle price will increase the net long straddle position; the net long futures position will increase (decrease) if the agent is long (short) in open futures.

In the model reformulated in terms of open futures and straddles, the comparative statics of a change in futures price with the straddle premium held constant illustrate the adjustments when the producer differs from the market in terms of his/her perception of the expected price but not on the price volatility. This allows us to resolve the ambiguity arising from results 3 and 4. Because of our centering $k = \bar{p}$ and of the assumption of symmetry in the distribution of \bar{p} , a *ceteris paribus* change in f will not affect the straddle price t (although the price of puts and calls will be affected through their delta). In this light, the own-price effects on the demand for futures of result 7 are intuitively appealing. For example, starting at the unbiased solution $q^* = s^* = 0$, a decrease in f other things being equal (i.e., an increase in the agent's expected price relative to the market) leads to a net long position.

Similarly, the comparative statics of a change in t with f held constant illustrates the adjustments when the producer differs from the market in terms of his/her perception of the volatility of price but not on the expected price. For example, a *ceteris paribus* decrease of the straddle price t is equivalent to the producer perceiving an increase in the volatility of the futures price \bar{p} relative to the market. If one evaluates the comparative statics effects at the unbiased point $q^* = s^* = 0$, the own-price effect of straddle demand implies a movement toward a long straddle, a strategy that is useful when the market is understating the volatility of \bar{p} .

The cross-price effects can be understood by considering the underlying use of straddles as hedging instruments to offset the futures position. Thus, in the case of a decrease in f , the

agent will change his/her straddle position to hedge against the net change in futures position. For $q > 0$, this means an increase in s . For $q < 0$, this means that the absolute value of q decreases so that fewer straddles are needed to hedge the position.

Pitfalls of Mean-Variance Analysis

The introduction pointed out that the truncation and nonlinearity introduced by options leads to a violation of commonly used justifications for mean-variance analysis. The results of the preceding section gives us an opportunity to illustrate this point further. Consider a mean-variance formulation of this model. For the profit equation (14) we have

$$(22.1)$$

$$E(\tilde{\pi}) = fy - c(y) + (\bar{p} - f)q + (\bar{w} - t)s, \text{ and}$$

$$(22.2) \quad \text{Var}(\tilde{\pi}) = q^2 \text{Var}(\tilde{e}) + s^2 \text{Var}(|\tilde{e}|)$$

because $\text{Cov}(\tilde{e}, |\tilde{e}|) = 0$ due to the assumed symmetry in $g(e)$.

Maximizing any function that is increasing in $E(\tilde{\pi})$ and decreasing in $\text{Var}(\tilde{\pi})$, as given in (22), would imply $s^* = 0$ for $\bar{w} = t$. Hence, in the mean-variance framework the optimal straddle position is zero if the straddle price is unbiased, regardless of the optimal position in the futures. This conclusion contrasts sharply with the results of the previous section. Because the distribution of profit cannot be normal in the presence of options, the ultimate rationale for mean-variance analysis must hinge on the undesirable assumption of quadratic utility.¹⁴

Conclusions

In this paper, we have analyzed the production and hedging decisions of the competitive firm facing both futures and options markets within an expected utility model that allows for basis risk. When futures prices and options premiums are perceived as unbiased, options are redundant hedging instruments. The optimal hedging strategy involves using only futures, and the amount of futures is determined by the covariance of cash and futures prices. However, if futures prices and/or options premiums are perceived as biased,

¹³ Proof of this result is available from the authors upon request.

¹⁴ Note that this utility function displays increasing absolute risk aversion with $u''' = 0$, which violates the conditions used to derive result 6 and rules out preference for positive skewness.

options are typically used along with futures. Thus, in this model options are appealing more as speculative tools to exploit private information on the price distribution, and less so as an alternative hedging instrument.

The qualitative effects of biased prices on the use of options and futures was investigated in terms of comparative statics effects. The nature of the speculative activity brought about by biased prices is clarified if the problem is formulated in terms of open futures and straddle positions. The sign of the open position in the futures depends only on the bias in the futures price, and (long) straddles are used even if the straddle price is unbiased whenever the futures position is open. In this context, options emerge as a useful device to insure against the price risk in an open speculative position.

[Received May 1989; final revision received April 1990.]

References

- Antonovitz, F., and R. D. Nelson. "Forward and Futures Markets and the Competitive Firm Under Price Uncertainty." *S. Econ. J.* 55(1988):182-95.
- Baron, D. P. "Price Uncertainty, Utility, and Industry Equilibrium in Pure Competition." *Int. Econ. Rev.* 11(1970):463-80.
- Batlin, C. A. "Production under Price Uncertainty with Imperfect Time Hedging Opportunities in Futures Markets." *S. Econ. J.* 49(1983):681-92.
- Benninga, S., R. Eldor, and I. Zilcha. "The Optimal Hedge Ratio in Unbiased Futures Markets." *J. Futures Mkts.* 4(1984):155-59.
- Black, F., and M. Scholes. "The Pricing of Options and Corporate Liabilities." *J. Polit. Econ.* 81(1973):637-54.
- Chavas, J.-P., and R. Pope. "Hedging and Production Decisions Under a Linear Mean-Variance Preference Function." *West. J. Agr. Econ.* 7(1982):99-110.
- Cox, J., and M. Rubinstein. *Options Markets*. Englewood Cliffs NJ: Prentice-Hall, 1985.
- Danthine, J. P. "Information, Futures Prices, and Stabilizing Speculation." *J. Econ. Theory* 17(1978):79-98.
- Feder, G., R. E. Just, and A. Schmitz. "Futures Markets and the Theory of the Firm under Price Uncertainty." *Quart. J. Econ.* 94(1980):317-28.
- Grant, D. "Theory of the Firm with Joint Price and Output Risk and a Forward Market." *Amer. J. Agr. Econ.* 67(1985):630-35.
- Holthausen, D. M. "Hedging and the Competitive Firm under Price Uncertainty." *Amer. Econ. Rev.* 69(1979):989-95.
- Honda, Y. "Production Uncertainty and the Input Decision of the Competitive Firm Facing the Futures Market." *Econ. Letters* 11(1983):87-92.
- Ishii, Y. "On the Competitive Firm under Price Uncertainty: Note." *Amer. Econ. Rev.* 67(1977):768-69.
- Kahl, K. H. "Determination of the Recommended Hedging Ratio." *Amer. J. Agr. Econ.* 65(1983):603-5.
- Karp, L. S. "Methods for Selecting the Optimal Dynamic Hedge When Production Is Stochastic." *Amer. J. Agr. Econ.* 69(1987):647-57.
- Losq, E. "Hedging with Price and Output Uncertainty." *Econ. Letters* 10(1982):65-70.
- Meyer, J. "Two-Moment Decision Models and Expected Utility Maximization." *Amer. Econ. Rev.* 77(1987):421-30.
- Newbery, D. M. G., and J. E. Stiglitz. *The Theory of Commodity Price Stabilization*. Oxford: Clarendon Press, 1981.
- Paroush, J., and A. Wolf. "Production and Hedging Decisions in Futures and Forward Markets." *Econ. Letters* 21(1986):139-43.
- Pratt, J. W. "Risk Aversion in the Small and in the Large." *Econometrica* 32(1964):122-36.
- Robison, L., and P. Barry. *The Competitive Firm's Response to Risk*. New York: Macmillan Co., 1987.
- Rolfo, J. "Optimal Hedging under Price and Quantity Uncertainty: the Case of a Cocoa Producer." *J. Polit. Econ.* 88(1980):100-116.
- Sandmo, A. "On the Theory of the Competitive Firm Under Price Uncertainty." *Amer. Econ. Rev.* 61(1971):65-73.
- Tsiang, S. C. "The Rationale of the Mean-Standard Deviation Analysis, Skewness Preference, and the Demand for Money." *Amer. Econ. Rev.* 62(1972):354-71.
- Wolf, A. "Optimal Hedging with Futures Options." *J. Econ. and Bus.* 39(1987):141-58.

Dynamically Optimal After-Tax Grain Storage, Cash Grain Sale, and Hedging Strategies

Russell Tronstad and C. Robert Taylor

This article utilizes a stochastic dynamic programming (SDP) model that considers the state variables of (a) before-tax income, (b) grain storage, (c) quantity of futures position, (d) value of futures position, (e) wheat price, and (f) basis level. Decision variables are monthly cash grain sales and futures market transactions. In comparing the post-sample performance of SDP to other marketing strategies over a four-year period, SDP resulted in \$5,961 to \$25,021 more wealth than the other strategies considered. Also, these other strategies yielded a standard deviation of after-tax income that was 30% to 621% greater than that from the SDP framework.

Key words: before-tax income, certainty equivalence, marketing strategies, stochastic dynamic programming.

Many studies have examined the optimal amount of grain that producers should sell or hedge via cash and futures markets under varying assumptions of market position, futures and cash price determination, unbiasedness of futures prices, production, utility function, and risk aversion (e.g., Berck, Heifner, Johnson, Kahl, Karp, Peck, Rolfo, Ward and Fletcher). However, these studies have failed to recognize many factors faced by producers in marketing grain. For example, producers will have different income levels and thus income tax rates as the price of grain varies. Thus, under a cash accounting system, the current income of producers will affect whether they should store grain into the next tax period, sell in the futures market, or sell in the cash market. In analyzing the effects of the 1986 Tax Reform Act on optimal cash grain marketing decisions, Tronstad and Taylor (1989b) solved for dynamically optimal after-tax marketing decisions utilizing a stochastic dynamic programming (SDP) model. They found that the sensitivity of optimal cash grain sales to income tax considerations could be magnified by the 1986 Tax Reform Act (TRA) because of larger jumps in marginal tax rate schedules. Thus, income tax considerations are likely important in determining optimal hedging decisions, even though the

1986 TRA decreased the marginal tax rate for most individuals. Also, the effects of government programs are not incorporated into the optimal hedging decisions of the above studies, and most grain producers participate in government programs.

The purposes of this study are to (a) extend the after-tax cash grain-marketing model of Tronstad and Taylor (1989b) to include after-tax hedging decisions; and (b) explore the benefits of an optimal dynamic marketing decision rule for grain producers which incorporates the discrete aspects of hedging decisions (i.e., progressive income tax structure and lumpiness of futures contracts), the before-tax income level of grain producers, the stochastic nature of cash price and basis levels without assuming certainty equivalence (i.e., with respect to the nonlinearities of the tax function), and government payments. As in Tronstad and Taylor (1989b), a representative Montana winter wheat producer was used to illustrate how these factors affect grain-marketing decisions and firm profitability because a much simpler model formulation results for a Montana grain producer than a grain producer from a multiple-crop region. That is, as a result of moisture limitations, Montana dryland grain producers essentially have only winter wheat as an economically viable crop alternative. However, the single-crop model can give insight into optimal decisions for a multiple-crop producer because the focus is on method and

Russell Tronstad is an extension economist, University of Arizona; C. Robert Taylor is Alfa Professor of Agricultural Policy, Auburn University.

timing of grain sales rather than crop selection. Forward contracting and options are not considered because futures contracts are similar to forward contracting and the minimum selling price concept of options is partially captured by government loan and target prices. As in Tronstad and Taylor (1989b), optimal grain-marketing decisions are obtained for a multiperiod horizon where the producer is presumed to maximize expected wealth.

Two commonly utilized frameworks which incorporate time and uncertainty in decision problems are SDP and Ito's stochastic control theory. SDP is selected for this analysis because (a) the discreteness of income tax schedules and futures contracts is more amenable to an SDP than stochastic control theory formulation, (b) an explicit solution is often not achieved for many problems when using optimal control theory (Burt), and (c) direct enumeration techniques are the only alternative to SDP and these methods are computationally inefficient (Bellman).

The next section presents the SDP formulation of the grain-marketing model. Cash price and basis transition probabilities are discussed in the third section; marginal tax schedules and other critical features are delineated in the fourth section.¹ The fifth section presents slices of the optimal grain-marketing decision rule generated from the SDP model. Then, the post-sample performance of the SDP marketing approach is

compared to other cash and futures marketing strategies from August 1985 to January 1989. Finally, the usefulness of an SDP approach is discussed.

Stochastic Dynamic Programming Model

To keep the SDP model formulation manageable, only March futures contracts are considered. The fact that winter wheat is a storable commodity makes this restriction reasonable because market conditions which affect old (new) crop futures contracts will usually affect all other old (new) crop futures contracts in a similar manner. However, even if an inverse market exists across adjacent futures contract months, immediate cash sales would likely be preferred to selling in the futures market. Thus, any benefit from adding another futures horizon is probably quite small. March futures contracts are utilized so the producer will be allowed to (a) hedge grain into the next tax year, (b) hedge anticipated production before the main growing season of winter wheat, and (c) receive almost all of the carrying charge of storage from hedging before the new U.S. harvest begins, as illustrated by Working.

The decision maker is assumed to maximize the expected present value of after-tax wealth over a multiyear planning horizon subject to the state variables of monthly cash price, basis, before-tax income, grain storage, futures position, and value of March futures position. Formulation of this problem as a SDP model results in the following recursive equation:

¹ Basis is defined as futures minus the cash price in this analysis. This definition is frequently found in academic literature, whereas the trade literature usually defines basis as cash minus the futures price.

$$\begin{aligned}
 (1) \quad & V_t(P_t, B_t, I_t, ST_t, QM_t, VM_t) = \underset{XC_t, XM_t}{\text{MAX}} E\{T[R(P_t, B_t, I_t, ST_t, QM_t, VM_t)] \\
 & \quad + \beta \cdot V_{t+1}(P_{t+1}, B_{t+1}, I_{t+1}, ST_{t+1}, QM_{t+1}, VM_{t+1})\} \\
 & \text{subject to} \\
 (2) \quad & I_{t+1} = I_t + P_t \cdot XC_t + VI_t + DEF_t - C_t \\
 (3) \quad & I_t = 0; \text{ if } t = 1, 13, 25, \dots \\
 & \quad \text{or the month of January} \\
 (4) \quad & 0 \leq XC_t \leq ST_t + Q_t \\
 (5) \quad & 0 \leq ST_t \leq SC \\
 (6) \quad & ST_{t+1} = ST_t + Q_t - XC_t \\
 (7) \quad & QM_{t+1} = QM_t + XM_t \\
 (8) \quad & -XM_t = QM_t; \text{ if } t = 3, 15, 27, \dots, \\
 & \quad \text{or the month of March}
 \end{aligned}$$

Before-tax income constraints

 Grain sale constraint

 Storage capacity constraint

 Storage constraint

 Futures constraints

- | | |
|---|---|
| $\left. \begin{aligned} (9) \quad & -QM_t \leq XM_t \leq ST_t + AQ_t - QM_t \\ (10) \quad & 0 \leq QM_t \leq ST_t + AQ_t, \text{ if} \\ & \quad ST_t + AQ_t < SC; \text{ else} \\ & 0 \leq QM_t \leq SC \end{aligned} \right\}$ | Legitimate hedging constraints |
| $(11) \quad VM_{t+1} = VM_t + XM_t \cdot (P_t + B_t) - VI_t \quad \}$ | Value of futures constraint |
| $\left. \begin{aligned} (12) \quad & VI_t = [(P_t + B_t) - (VM_t/QM_t)] \cdot XM_t \\ & \quad \text{if } XM_t < 0, \text{ else} \\ & VI_t = 0 \end{aligned} \right\}$ | Loss (-) or profit (+) on short futures position closed |
| $(13) \quad P_{t+1} = f_1(P_t, e_{1t}) \quad \}$ | Stochastic Markovian price relationship |
| $(14) \quad B_{t+1} = f_2(B_t, e_{2t}) \quad \}$ | Stochastic Markovian basis relationship |

where t is the monthly index; $V_t(\cdot)$ is the maximum expected present value of after-tax income from month t through the multiyear terminal month T given the initial state; P_t is the mid-month price of Montana winter wheat; B_t is the mid-month Kansas City Board of Trade (KCBT) March futures price minus P_t ; I_t is the beginning before-tax income state of the producer in month t ; ST_t is the beginning storage state of the grain producer in month t ; QM_t is the beginning quantity of March futures position (a positive value indicates a short position); VM_t is the beginning value of March futures position (average transaction price (ATP_t) of a short futures position multiplied by the quantity short); XC_t is cash winter wheat sales; XM_t is March futures transactions [positive (negative) values denote selling (buying) of March futures contracts]; E is the expectation operator; $T\{R(\cdot)\}$ is after-tax income as a function of before-tax revenue, $R(\cdot)$; β is the real discount factor ($1/1.005$); Q_t is production for month t (nonzero only for the month of August); SC is total storage capacity; AQ_t is anticipated production (positive for the months of April through August); VI_t is the profit or loss generated from closing out a short futures position; DEF_t is deficiency payments; C_t is production, commission and margin expenses; $f_i(\cdot)$'s ($i = 1, 2$) are stochastic Markovian relationships with associated random variable e_{it} 's.

Before-tax income (I_t), cash grain sale (XC_t), storage (ST_t), and price (P_t) constraints in equations (2) through (6) and (13) are quite similar to, and the only constraints necessary for, an SDP model that considers cash grain sales alone as solved by Tronstad and Taylor (1989b). However, the primary difference occurs when before-tax income gained or lost from hedging (i.e., VI_t , commission and/or margin expenses) enters constraint (2). Hedging and basis level

constraints [i.e., equations (7)–(12) and (14)] influence VI_t and C_t in equation (2) indirectly.

I_t , ST_t , QM_t , and VM_t are deterministic state variables in the above formulation. Because of uncertainty associated with cash price (P_t) and basis (B_t) movements, P_t and B_t are treated as stochastic state variables. Even though P_t and B_t enter the before-tax function [$R(\cdot)$] linearly, certainty equivalence requirements are not satisfied for P_t or B_t because the after-tax function [$T\{\cdot\}$] is nonlinear (Simon, Theil). Subsequently, P_t and B_t must be treated as true stochastic state variables in the SDP model formulation.

Except for January, ending before-tax income (I_{t+1}) for each period is before-tax income carried over from the previous month (I_t) plus current cash grain sale receipts ($P_t \cdot XC_t$), hedging profits (positive VI_t), and deficiency payments (DEF_t); minus current costs of production, commission and margin expenses (C_t), and hedging losses incurred in month t . The producer is assumed to comply with government program requirements, as is the case for virtually all Montana dry-land grain producers. I_t begins each tax year (i.e., January) with zero before-tax income (constraint 3) because grain sales, profits or losses on hedging, deficiency payments, and before-tax income accumulated prior to January are accounted for by taxable income calculated in December of the previous year. Ending storage for month t (ST_{t+1}) is increased by grain harvested during August (Q_t) and decreased by cash grain sales (XC_t) as given in equation (6).

The ending month's quantity of March futures position (QM_{t+1}) is increased by selling March futures contracts (i.e., positive XM_t) and decreased by buying March futures contracts (i.e., negative XM_t). If a March futures position exists (i.e., $QM_t > 0$) at maturity, the futures position must be closed as indicated in equation (8). Equations (9) and (10), called legitimate hedg-

ing constraints, allow the producer to maintain a bona fide IRS hedging account. That is, the producer cannot have a long futures position ($0 \leq QM_t$) or a short futures position that is greater than current storage plus anticipated production ($QM_t \leq ST_t + AQ_t$) or storage capacity ($0 \leq QM_t \leq SC$). This constraint also implies that current futures transactions (XM_t) are bound between the negative of any current futures position ($-QM_t$) and grain in storage plus anticipated production minus any futures position ($ST_t + AQ_t - QM_t$). As indicated in (11), the ending monthly value of March futures contracts (VM_{t+1}) increases or decreases as a result of March futures transactions by the amount of $XM_t \cdot (P_t + B_t)$. Also, to maintain the original transaction price of March futures contracts sold when part

temporaneous basis because it should capture most of the information about recent transportation and market conditions between KCBT March futures price and P_t . Lagged cash prices are hypothesized to positively influence the basis because of increased storage costs. Maturity time influences the basis because storage costs are directly related to the length of time between the current month and the contract month. A negative time trend for the basis is anticipated because the Portland market has been stronger than the Gulf market in recent years.

Utilizing Schwarz's Bayesian information criteria (BIC)² (Schwarz) to determine appropriate lag lengths and variables, the following cash price and basis relationships were estimated (t -values are in parentheses):

$$\begin{aligned}
 (15) \quad P_t &= .4547 + .8960 \cdot P_{t-1} - .07815 \cdot HD_t \\
 &\quad (2.900) \quad (24.417) \quad (-2.406) \\
 &\quad - .00094026 \cdot MT_t + e_{1t} \quad \bar{R}^2 = .8982 \quad DH = -.1852 \\
 &\quad \quad \quad (-1.945) \\
 (16) \quad B_t &= .6185 \cdot B_{t-1} + .2548 \cdot B_{t-2} + .01359 \cdot TM_t + e_{2t} \quad \bar{R}^2 = .7906 \quad DH = -1.1848 \\
 &\quad (6.8753) \quad (3.0369) \quad (2.7761)
 \end{aligned}$$

of a short futures position is closed out, a profit (loss) of VI_t is subtracted (added) from (to) the amount associated with closing out part or all of a short futures position.

Cash Price and Basis Transition Probabilities

As in Tronstad and Taylor (1989b), Markovian transition probabilities for cash price were determined by estimating P_t as a function of own lagged price ($P_{t-j}; j = 1, 2, 3$), a harvest dummy variable (i.e., 1 for the months of June, July, and August; 0 otherwise), a time trend, and autoregressive error ($e_{t-j}; j = 1, 2, 3$). Lagged prices capture current grain market conditions. A harvest dummy variable for the months of June, July, and August captures any dampening effect on P_t of the harvest of hard red winter wheat. A time trend variable represents changes in P_t resulting from technological advances (e.g., new crop varieties, herbicides, and pesticides, in conjunction with no-till farming practices).

Markovian basis relationships were determined by estimating basis as a function of lagged basis ($B_{t-j}; j = 1, 2, 3$), lagged cash price levels ($P_{t-j}; j = 1, 2, 3$), time of basis to maturity (i.e., 0 to 11 months), and a time trend variable. Own-lagged basis is hypothesized to influence con-

where P_t is the mid-month Montana winter wheat price (\$/bu.) (*Agricultural Prices*), HD_t is a harvest dummy variable, MT_t is a monthly time trend, B_t is the basis level or mid-month KCBT March Futures (*Wall Street Journal*) minus P_t , TM_t is the number of months until maturity, e_{it} ($i = 1, 2$) is a normally distributed error term with a constant variance, \bar{R}^2 is the adjusted coefficient of determination, and DH is Durbin's H -statistic. The monthly data series runs from January 1977 to August 1988 and all prices are deflated by the index of prices paid by farmers (*Agricultural Prices*), utilizing August 1988 as a base period. Also, to reduce model complexity, the second-order Markov process of B_t in equation (16) was reduced to a first-order Markov process as described in Burt and Taylor.

Marginal Tax Schedules and Other Critical Input Features

The 1988 federal income tax code has three marginal tax rates. Marginal tax rates for married individuals filing jointly are 15% for tax-

² $BIC(j)$ chooses the j th model such that $2\ln(M_j) - k\ln(n)$ is largest, where M is the maximum likelihood function, k is the number of parameters associated with the j th model, and n is the number of observations.

able incomes less than \$29,750, 28% for taxable income between \$29,750 and \$71,901, 33% for taxable income between \$71,901 and \$171,090, and 28% for taxable income exceeding \$171,090. Standard deductions for a married couple (i.e., \$5,000) and two personal exemptions (i.e., \$3,900) are claimed in determining federal income tax liabilities.

State taxes for Montana are made up of ten relatively flat marginal tax brackets. After a \$4,280 standard deduction, the marginal tax rate is 2% for income less than \$1,400, 3% for income between \$1,400 and \$2,900, 4% for income between \$2,900 and \$5,700, 5% for income between \$5,700 and \$8,600, 6% for income between \$8,600 and \$11,400, 7% for income between \$11,400 and \$14,300, 8% for income between \$14,300 and \$20,000, 9% for income between \$20,000 and \$28,000, 10% for income between \$28,600 and \$50,000, and 11% for income greater than \$50,000. Also, social security taxes are included at a rate of 13.02% of income, with a maximum social security tax of \$5,702.80.

Grid dimensions utilized in this analysis are as follows; eight cash price (P_t) states (\$2.25, \$2.50, \$2.75, \$3.00, \$3.25, \$3.50, \$4.25, and \$5.00 per bushel); seven basis (B_t) states (-\$.75, -\$.15, \$.10, \$.35, \$.60, \$.85, and \$1.45 per bushel); seven income (I_t) states (-\$60,000 to \$60,000 in \$20,000 increments); nine storage (ST_t) states (0 bu. to 40,000 bu. in 5,000 bu. increments); nine quantity of March futures (QM_t) states (0 bu. to 40,000 bu. in 5,000 bu. increments); and eleven value of March futures (VM_t) states (\$1.50/bu. to \$6.50/bu. in \$.50 increments, multiplied by QM_t). The lower range of the state space for P_t (\$2.25/bu.) is essentially the current CCC loan rate of \$2.21 per bushel (\$2.28 in 1987). The distribution of P_t is somewhat skewed toward higher prices because of the CCC loan rate. Therefore, to reduce computation requirements, the upper end of the distribution of P_t is given a coarser grid space. Because the distribution of B_t is essentially symmetric, both tails of the distribution of B_t are given a coarser grid space.

The mid-point between each cash price and basis state increment is utilized in conjunction with equations (15) and (16) in calculating Markovian transition probabilities. For example, if the current cash price state is \$3.00 per bushel, equation (15) is utilized to determine an expected price for $t + 1$ in t so that the probability of cash going to a price state of \$3.25 per bushel was determined by integrating the area of cash

price falling between \$3.375 per bushel and \$3.125 per bushel in period $t + 1$, given that the error structure of equation (15) is normally distributed. This process was done for all cash price and basis states so that the product of the two probability matrices for each period would sum to one in going from one $\{P_t, B_t\}$ state combination to all possible $\{P_{t+1}, B_{t+1}\}$ combinations. These transition probabilities are utilized to calculate optimal expected values from one period to the next [i.e., $V_t(\cdot)$ from $V_{t+1}(\cdot)$] and determine an optimal converged decision rule. The reader is referred to Howard's book, *Dynamic Programming and Markov Processes*, for a more thorough discussion of this process.

Because before-tax income is always zero at the beginning of every tax year, -\$60,000 is about the lower limit of before-tax income that the firm could attain by the end of the tax year given production expenses, even if zero cash grain sales occurred. Conversely, with production expenses and "normal" production levels, the firm is unlikely to achieve a before-tax income level above \$60,000. Thus, I_t was discretized in increments of \$20,000 from -\$60,000 to \$60,000. ST_t , QM_t , XC_t , and XM_t were discretized in increments of 5,000 bushels because KCBT futures contracts are traded in 5,000 bushel increments.

The above state space adequately represents the state space of the producer, in that the unconditional probability of values occurring outside the state space is small. Discretization of continuous state variables is based on a subjective trade-off between computation requirements and sensitivity of the optimal decision rule to changes in the level of state variables. However, to minimize any bias associated with discretization of continuous state variables, linear interpolation of the optimal value function [i.e., values of V_{t+1} in the right-hand side of equation (1)] was utilized whenever the resulting state values in $t + 1$ did not exactly match the state space described above.

Critical production characteristics for this representative firm are farm payment acreage and crop production acreage of 1,000 acres; variable per acre production costs incurred in May (spraying), August (harvest), and September (planting) of \$6, \$6, and \$36.32, respectively; overhead costs of \$35,800/12 are charged for each month; and storage capacity (SC) of 40,000 bushels. Production costs are primarily based on Economic Research Service production costs for hard red winter wheat in the Northern Plains. Commission costs of \$37.50 are incurred for each

one-way trade of a 5,000 bushel futures contract. An initial margin of \$1,500 is required for each futures contract. Margin money requirements increase or decrease with futures price movements in order to maintain the initial margin cushion of \$1,500 per contract, or \$.30 per bushel. Thus, the margin expense or revenue incurred each month is the interest paid (0.5% each month) on the net balance in the producer's margin account. Other than the margin expense or revenue realized each month, the loss or profit on futures price movements does not accrue to before-tax income until a futures position is closed out.

Yield is estimated as a function of a time trend using annual Montana winter wheat yields from 1962 to 1987 (period of available data). Because 1985 was an unusually dry year compared to the other years in this period, a dummy variable was included for 1985. The following equation was estimated for Montana winter wheat yields (*t*-values are given in parentheses)

$$(17) \quad Y_t = 26.985 + .2361 \cdot AT_t - 16.42 \cdot D85, \\ (19.437) \quad (2.460) \quad (-4.646) \\ \bar{R}^2 = .472 \quad D-W = 1.533,$$

where Y_t is yield (bu./acre) (*Crop Production*), AT_t is an annual time trend variable (1962 = 1), $D85$ is a dummy variable for 1985 (1 if 1985, 0 otherwise), \bar{R}^2 is the adjusted coefficient of determination, and $D-W$ is the Durbin-Watson statistic. Y_t is utilized to determine production (Q_t) and anticipated production (AQ_t) in the SDP model formulation. Thus, both Q_t and AQ_t are simply Y_t multiplied by 1,000 acres.

Deficiency payments are distributed equally throughout the year as follows:

$$(18) \quad DEF_t = (TP_t - P_t) \cdot FPP/12; \\ \text{if } P_t < TP_t, \text{ else} \\ DEF_t = 0$$

where DEF_t is the deficiency payment received by the producer, TP_t is the target price (\$4.38, \$4.23, \$4.10, and \$4.00, prior to 1987, 1988, 1989, 1990 and beyond crop marketing years, respectively), and FPP is farm program production, set at 30,000 bushels. Equation (18) is a proxy to actual deficiency payments because P_t is used to approximate the average national winter wheat price. Also, caution should be exercised in using DEF_t when prices are above and below TP_t for the crop marketing year because the actual annual payment would be somewhat less than the cumulative monthly payment in (18).

Montana has a relatively dry and cool climate, and the loss in grain weight from reduced moisture content and/or spoilage is considered negligible. Also, the only cost associated with on-farm grain storage in this analysis is the time value of money associated with grain storage. Commercial grain storage was not considered.

Inventory financing costs of storage are captured in the SDP model through the discount factor. Even though the nominal interest rate for borrowed funds is undoubtedly greater than it would be for loanable funds, the after-tax differential in real terms is probably quite small. Also, off-farm income, additional enterprises, and income tax management tools other than grain storage are not considered.

Optimal Cash Grain Sale, Storage, and Hedging Decisions

This section graphically presents optimal grain-marketing decisions derived from the SDP model. Optimal grain-marketing decisions are presented for April and December to illustrate the converged optimal grain-marketing decision rule generated for every month and state.³ April illustrates hedging decisions in relation to anticipated production, and December illustrates the effects of income tax considerations on optimal cash grain sale and futures transactions.⁴

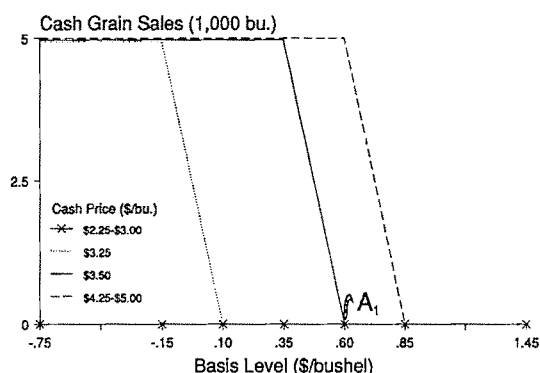
Figure 1 presents optimal marketing decisions for April where the state variables of cash price, basis, and storage vary. Because any short futures position (QM_t) in March must be closed, April always begins with no futures position. Thus, QM_t equals zero in figure 1, and QM_t equal to zero also implies that VM_t equals zero. The state variable of before-tax income (I_t) is also fixed at \$0.00 in figure 1 because income tax effects are less pronounced early in the tax year compared to months later in the tax year.

In interpreting figure 1, legend lines of cash price in conjunction with the basis level given on the horizontal axis determine the optimal level of cash grain sales and futures transactions. For example, if the current levels of cash price (P_t),

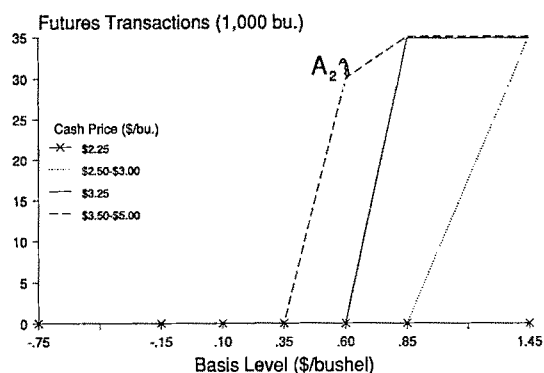
³ Convergence of the decision rule was obtained in 60 states, or 5 years.

⁴ Marketing decisions for additional months and scenarios are available from the authors upon request. Computation requirements for the model were 166 minutes of CPU time on a CRAY X/MP-48, or about 100 times greater than the computation requirements for just a cash grain sale model as done by Tronstad and Taylor (1989b). Thus, the incorporation of futures transactions into the grain-marketing decision greatly increases model complexity.

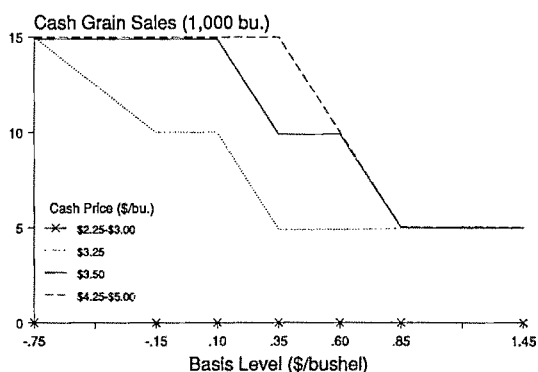
PANEL A. Storage Level = 5,000 bushels



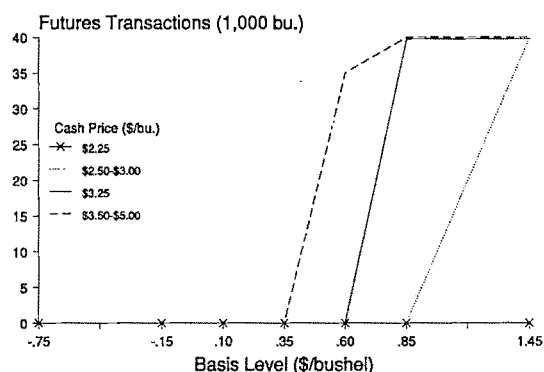
PANEL D. Storage Level = 5,000 bushels



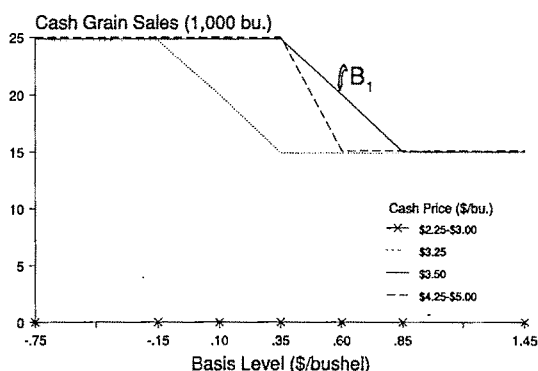
PANEL B. Storage Level = 15,000 bushels



PANEL E. Storage Level = 15,000 bushels



PANEL C. Storage Level = 25,000 bushels



PANEL F. Storage Level = 25,000 bushels

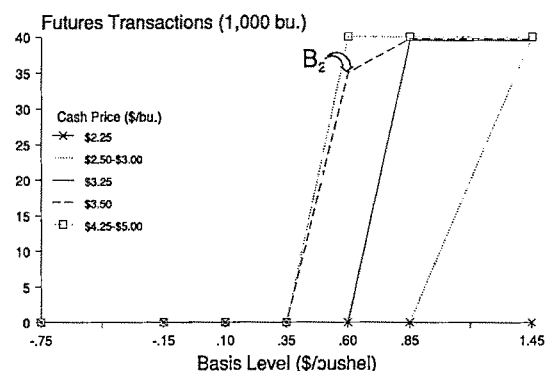


Figure 1. Optimal cash grain sales and futures transactions for April given a before-tax income level (I_t) = \$0.00, and quantity of March futures position (QM_t) = 0 bushels

basis (B_t), storage (ST_t), before-tax income (I_t), and short futures position (QM_t) for the firm are \$3.50 per bushel, \$.60 per bushel, 5,000 bushels, \$0.00, and 0 bushels, respectively, Panels A and D indicate that the firm should incur 0 bushels of cash grain sales (point A_1) and sell 30,000 bushels of March futures contracts (point

A_2). However, if the storage level for the firm were 25,000 bushels instead of 5,000 bushels, Panels C and F indicate that the firm should incur 20,000 bushels of cash grain sales (point B_1) and sell 35,000 bushels of March futures contracts (point B_2). Since optimal marketing decisions are linearly interpolated and occur in

5,000 bushels increments, cash price line "kinks" often appear in figure 1 at points where the basis level was discretized, as described earlier.

Results in figure 1 indicate that cash grain sales are generally preferred to hedging if the basis is less than \$.35 per bushel. However, because the basis level in conjunction with the cash price determine whether cash grain sales, futures transactions, and/or storage is most profitable, no cash grain sales or positive futures transactions occur for the lowest cash price state of \$2.25 per bushel, even when the basis level is at its lowest (-\$.75/bu.) or highest price state (\$1.45/bu.). For cash prices between \$2.50 and \$3.00 per bushel, storage is always preferred to cash grain sales and positive futures transactions do not occur until the basis level is above \$.85 per bushel. All of storage goes to cash grain sales for basis levels less than \$.35 per bushel and cash price levels above \$3.50 per bushel, whereas futures transactions occur at their upper limit whenever the basis and cash price levels are above \$.60 and \$3.50 per bushel, respectively.

Because harvest will occur before the upcoming March futures contract matures, the SDP model allows the producer to hedge anticipated production (AQ_t) in April (i.e., discrete amounts of AQ_t from 0 to 30,000 in 5,000 bushel increments). Therefore, the opportunity cost of selling grain in storage on the cash market is less for April through August since AQ_t can be substituted for storage (ST_t) to satisfy equations (9) and (10). Consequently, cash grain sales and futures transactions are greater for April than preceding months of the current tax year, given an initial storage level. Also, the time to maturity (TM_t) for March futures contracts is further away for April (i.e., eleven months) than any other month. Thus, *ceteris paribus*, higher basis levels are required to signal positive futures transactions for April than months which have a shorter time to maturity because TM_t had a significant and positive coefficient in explaining the basis level. In general, the basis level must be about \$.50 to \$.75 greater for lower cash price states (less than \$3.00) and about \$.25 to \$.50 greater for higher cash price states (greater than \$3.50) before hedging occurs at the same level for the month of April compared to months preceding March that are in the current tax year.

Figure 2 graphically depicts optimal cash grain sales and futures transactions for December. Because the month of December is crucial for income tax considerations, the state variables of before-tax income (I_t), basis (B_t), and cash price (P_t) vary in this figure. That is, the legend lines

of cash price in conjunction with the before-tax income state of the firm on the horizontal axes in figure 2 determine the optimal level of cash grain sales and futures transactions. For example, if the current levels of cash price, before-tax income, basis, storage, and short futures position for the firm are \$2.50 per bushel, -\$40,000, -.35 per bushel, 25,000 bushels, and 0 bushels, respectively, Panels B and E indicate that the firm should incur 20,000 bushels of cash grain sales (point A_1) and not take any short futures position (point A_2).

Results in figure 2 demonstrate several factors in relation to the potential income tax liabilities of the firm. First, as seen in Panels A, B, and C, cash grain sales generally decrease as the before-tax income level of the firm increases. This is because the marginal tax rate for the firm increases as the before-tax income level of the firm increases. Thus, grain storage is more generally profitable for a higher before-tax income state than a lower before-tax income state. However, as shown in Panel A of figure 2, cash grain sales increase for cash price states above \$3.25 per bushel and before-tax income levels above \$40,000. This is because the basis level (-\$.15/bu.) is not favorable for hedging transactions and the increase in the marginal tax rate is relatively small on additional grain sold if the before-tax income level is above \$40,000. Consequently, futures transactions can decrease as the before-tax income level of the firm increases at low basis levels.

Second, cash grain sales are greater for lower cash price states than higher price states at before-tax income levels less than -\$20,000. This is because cash grain sales for the high cash price states push the producer into a higher marginal tax bracket than the low cash price states. Also, cash grain sales are greater for December than preceding months at low cash price and before-tax income levels in order to take advantage of zero or near zero marginal tax rates. That is, for months preceding December, the cash price might increase before the next tax year begins, whereas December is the last month of opportunity for selling grain in the current tax year.

Although the effects of a March futures position (VM_t) or average transaction price (ATP_t) on optimal grain-marketing decisions are not graphically illustrated here, some general insights are noted. Results revealed that low ATP_t s usually have larger cash grain sales than high ATP_t s, because the revenue generated from closing a short futures position is greater for high ATP_t s than low ATP_t s. This additional revenue

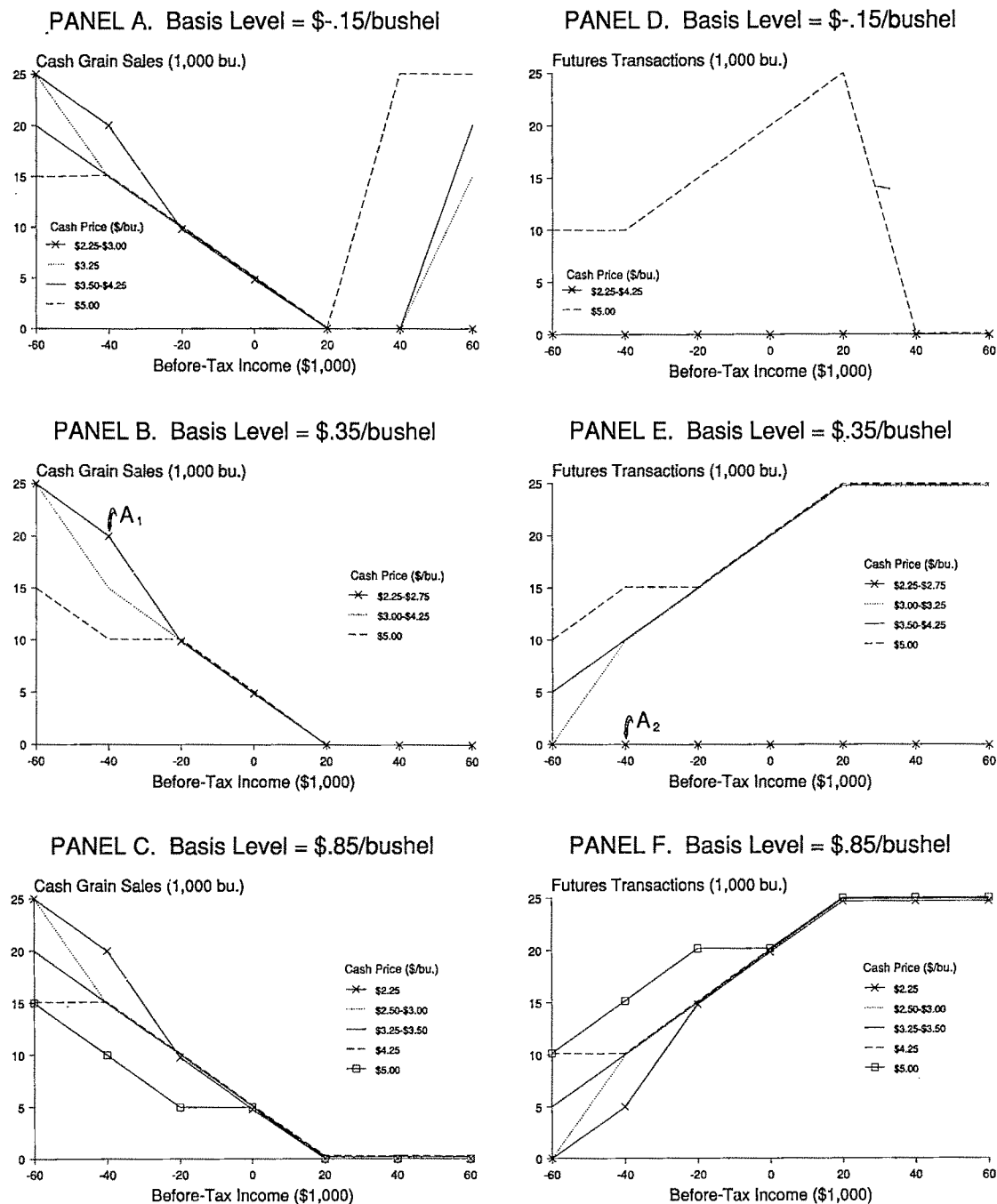


Figure 2. Optimal cash grain sales and futures transactions for December given a storage level (ST_t) = 25,000 bushels, and quantity of March futures position (QM_t) = 0 bushels

tends to increase the marginal tax rate of the firm and decrease the after-tax profitability of additional cash grain sales. Thus, lower optimal cash grain sales are generally given for higher ATP_t states.

The value associated with any short futures

position tends to be dominated by basis, cash price, and before-tax income level effects when determining optimal futures transactions. However, results suggest that lower ATP_t s are less profitable for closing out part or all of a short futures position for lower before-tax income

levels. This is because less revenue is generated from closing out a short futures position with a low *ATP*, than a high *ATP*, and income tax effects are not so dominating over the *ATP*, effect at lower before-tax income states.

In summary, results from the SDP grain model suggest that, not surprisingly, producers need to evaluate potential income tax liabilities as well as current market conditions when making grain-marketing decisions. To quantitatively evaluate the potential benefits from utilizing an SDP approach for grain-marketing decisions compared to other grain marketing strategies, the next section compares the post-sample performance of an SDP marketing strategy to other more traditional marketing strategies.

Comparison of Stochastic Dynamic Programming to Other Grain-Marketing Strategies

Profits are generated from August 1985 to January 1989, depending on the initial before-tax income level and grain-marketing strategy followed.⁵ Grain-marketing strategies considered were (a) 100% cash grain marketing with actual production divided equally for all months after harvest, (b) 100% utilization of futures contracts, (c) a minimum variance hedging ratio (MVHR) framework, and (d) the SDP model previously presented.

A MVHR represents the futures position that minimizes the variance associated with unequal price changes in the cash and futures markets during the hedge duration for a given case position (Witt, Schroeder, Hayenga). An MVHR framework is implementable since it does depend on individual preferences (Adler and Demple), and the MVHR is equivalent to an optimal hedging ratio under certain conditions (Heifner, Bond and Thompson). Utilization of a 100% cash or 100% futures marketing strategy would be equivalent to an MVHR strategy if the MVHR was 0 or 1, respectively. MVHRs were obtained by regressing annual cash price changes on futures price changes (e.g., see Witt, Schroeder, and Hayenga).

Futures transactions were closed during the contract maturity month of March. Thus, in calculating MVHRs for the post-sample evaluation, an additional observation was available

every March so that the MVHR was recalculated every year at this time. This procedure resulted in MVHRs of .6739, .6506, .8640, and .8778 for the time periods of March 1977 to March 1985, 1986, 1987, and 1988, respectively.⁶ For the SDP framework, cash price and basis transition probabilities were reestimated every year at harvest.⁷ Grain-marketing decisions from following either an all cash, all futures, MVHR, or SDP framework are illustrated in table 1.

Because the SDP model considers income tax liabilities, different grain-marketing decisions were generated depending on the initial before-tax income level of the firm in August 1985. Thus, different initial before-tax incomes are given in table 1 to illustrate income tax effects on grain-marketing decisions. Because most of the production costs occur after harvest, starting before-tax income states less than -\$20,000 were not considered.

For the 1985 tax year, starting with a before-tax income less than zero in August resulted in

⁶ In translating the continuous-static grain-marketing information of an MVHR into the discrete-dynamic nature of the hedging/production problem, the following steps were taken. First, actual production at the 1985 harvest was divided evenly between August 1985 and March 1986 (maturity time). Production allocated for each month was multiplied by the hedge ratio for the associated time period. A cumulative sum of the hedged amount for each month was started in August; and, when the cumulative sum exceeded 2,500 bushels, one March futures contract (5,000 bushels) was sold for that month. Similarly, when the cumulative sum exceeded 7,500 bushels, a second March futures contract was sold, and so on. After the quantity of grain to be hedged for this period was determined, remaining grain was evenly distributed by cash grain sales between August 1985 and March 1986. Then, all futures contracts that were previously sold were bought back in March and this grain was sold in the cash market. Results were essentially the same when using the month of February instead of March to close out futures contracts.

After March 1986, anticipated production was also utilized for making hedging decisions until harvest occurred. After harvest, the amount of grain hedged on anticipated production was subtracted from actual production. Then, this remaining grain was allocated between March futures contracts and cash grain sales as described above. The above grain-marketing procedure was utilized for all subsequent years in determining the grain-marketing activity generated by utilizing an MVHR framework or 100% utilization of futures.

The above procedure does not represent true MVHR. However, marketing decisions were also calculated from MVHRs that were updated monthly utilizing monthly differences of March futures prices. That is, in April the producer would make the initial hedge on anticipated production, and in August the producer's hedge would be based on actual production. Futures transactions were rounded to 5,000-bushel increments. Because the results of this procedure were inferior to the first procedure (i.e., on average, the cash position was \$9,924 less and the standard deviation of after-tax income was \$5,476 more as presented in table 3), these results are not reported here.

⁷ Equations estimated with data through August 1985, 1986, and 1987 are quite similar to those with data estimated through August 1988, i.e., equations (15) and (16). All equations are given in Tronstad and Taylor (1989a).

⁵ Expenses and revenues occur as described for the SDP model. Thus, the cash position is compounded monthly at a before-tax income rate of .5%.

Table 1. Harvest, Storage Level, and Short Futures Position of the Firm Illustrated Following Either an All-Cash, All-Futures, Minimum Variance Hedging Ratio (MVHR) or Stochastic Dynamic Programming (SDP) Strategy

Month	Harvest (thou. bu.)	Storage Level					All Futures	MVHR	SDP (-20 to 20)	SDP (40 to 60)	
		All Cash	All Futures	MVHR	SDP (-20)*	SDP (20)					SDP (60)
		(bu.)					(thou. bu.)				
8/85	16	14,667	16,000	15,250	16,000	16,000	16,000	0	0	0	0
9/85		13,333	16,000	14,500	16,000	16,000	16,000	5	5	0	0
10/85		12,000	16,000	13,750	16,000	16,000	16,000	5	5	0	0
11/85		10,667	16,000	13,000	0	6,000	11,000	10	5	0	0
12/85		9,333	16,000	12,250	0	1,000	11,000	10	5	0	0
1/86		8,000	16,000	11,500	0	1,000	6,000	10	10	0	0
2/86		6,667	16,000	10,750	0	1,000	6,000	15	10	0	0
3/86		5,333	0	0	0	1,000	6,000	0	0	0	0
4/86		4,000	0	0	0	1,000	6,000	5	0	0	0
5/86		2,667	0	0	0	1,000	6,000	5	5	0	0
6/86		1,333	0	0	0	1,000	6,000	10	5	0	0
7/86		0	0	0	0	1,000	6,000	10	5	0	0
8/86	32	29,333	32,000	30,500	32,000	33,000	38,000	15	5	0	0
9/86		26,667	32,000	29,000	32,000	33,000	38,000	15	10	0	0
10/86		24,000	32,000	27,500	32,000	33,000	38,000	20	10	0	0
11/86		21,333	32,000	26,000	32,000	33,000	38,000	20	15	0	0
12/86		18,667	32,000	24,500	7,000	8,000	23,000	25	15	0	0
:	:	:	:	:	:	:	:	:	:	:	:
8/88	19	17,417	19,000	18,500	24,000	24,000	29,000	10	10	20	25
9/88		15,833	19,000	18,000	24,000	24,000	29,000	10	10	20	25
10/88		14,250	19,000	17,500	24,000	24,000	29,000	15	15	20	25
11/88		12,667	19,000	17,000	24,000	24,000	29,000	15	15	20	25
12/88		11,083	19,000	16,500	19,000	19,000	29,000	15	15	15	25
1/89		0	0	0	0	0	0	0	0	0	0

^a Numbers in parentheses refer to the initial before-tax income level, e.g., (-20) refers to -\$20,000.

a strategy where all of the grain that was harvested should be sold in the current tax year. However, starting with a before-tax income of \$60,000 yielded a decision of selling only 5,000 bushels in the cash market for the current tax year. This demonstrates the importance of selling grain in storage to capitalize on low marginal tax rates and maintaining storage at relatively high marginal tax rates to reduce income tax liabilities. Winter wheat cash price levels were relatively low throughout 1986 so that cash grain sales did not occur in the SDP marketing strategy until December. Thus, even though the current cash price was relatively low, the after-tax profitability of selling some grain at a low marginal tax rate resulted in positive cash grain sales for December 1986. Cash prices continued to be relatively low throughout 1987, so the storage capacity of the firm (40,000 bu.) was reached in August 1987 with a harvest of 36,000 bushels. Thus, depending on the initial before-tax income for the SDP strategy, forced cash grain sales between 3,000 and 19,000 bushels occurred in August 1987. Cash price and basis levels moved upward through the latter part of 1987, and March futures contracts were sold in

October, November, and December utilizing the SDP framework.

As a result of widespread low soil moisture conditions in May of 1988, grain prices moved upward. This increase in prices indicated that March futures contracts and cash grain sales should occur under the SDP strategy. However, because of the severity of the drought, cash and futures price levels continued to rise throughout the summer of 1988. Because actual production at harvest was substantially less than that anticipated, the firm was forced to buy back some March futures contracts in August so that the short futures position of the firm would not exceed grain storage. Buying back March futures contracts at a much higher price in August and later on (including the termination point) resulted in a substantial before-tax income loss (anywhere from \$13,224 to \$23,948) for the SDP framework.

Also, the SDP decision rule with the highest initial before-tax income had the most grain in storage carried over every tax year because the benefit from storing grain into the following tax year is greater at high than low before-tax income states due to the progressive marginal tax

structure. Similarly, the short futures position of the firm was greater for the high than low initial before-tax incomes because cash grain sales are less lucrative for the high than low initial before-tax income levels.

Table 2 illustrates the cash position of the firm for the marketing strategies reported in table 1. Throughout most of the period, the cash position of the firm for the SDP framework is below (above) other grain-marketing strategies considered if the initial before-tax income was relatively high (low). This is because grain storage or less cash grain sales are desirable to reduce the income tax liabilities of the firm given a high initial before-tax income level. Conversely, cash grain sales are more desirable at lower initial before-tax income levels.

Table 3 analyzes the above marketing strategies by giving the mean and standard deviation of after-tax profits for the August 1985 through December 1988 period, the before-tax termination money received in January 1989 from closing out any outstanding futures position and selling all grain in storage on the cash market, the cash position of the firm at January 1989. The table also indicates the marginal tax rate that

would be required on all the excess before-tax termination income received in January from the SDP strategy over other strategies so that equal after-tax terminal wealth values would exist in January 1989. Thus, marginal tax rate figures for equality are calculated by dividing the differential in cash position on January 1989 with the differential of before-tax termination income received in January 1989, given that the SDP strategy received a higher before-tax termination value in January than the alternative strategy compared.

Results in table 3 show that the mean of after-tax income for the SDP strategy in the four tax years considered is sometimes less than that produced from the other marketing strategies. However, because of the additional revenue realized from selling all grain in storage and closing all futures positions at January 1989, the cash position of the firm at the termination point is anywhere from \$5,961 to \$25,021 higher for the SDP decision rule than all of the other marketing strategies considered. Also, 32% is the lowest marginal tax rate that would be required on all of the additional before-tax income received from the SDP strategy over all other strategies

Table 2. Cash Position of the Firm Illustrated Depending on the Initial Before-Tax Income Level and the Marketing Strategy Followed

Month:	Initial Before-Tax Income = -\$20,000				Initial Before-Tax Income = \$60,000			
	All Cash	All Futures	MVHR ^a	SDP	All Cash	All Futures	MVHR	SDP
8/85	-24,883	-29,083	-26,721	-29,083	55,517	51,317	53,679	51,317
9/85	-59,978	-68,570	-63,758	-68,532	20,824	12,232	17,044	12,270
10/85	-58,861	-71,893	-64,582	-71,858	22,345	9,313	16,624	9,348
11/85	-57,472	-75,276	-65,266	-19,201	24,140	6,336	16,346	23,911
12/85	-43,116	-65,676	-52,994	-9,320	27,000	12,698	20,849	23,986
1/86	-41,715	-68,981	-53,693	-12,350	28,751	9,785	20,520	38,373
2/86	-40,574	-72,337	-54,500	-15,395	30,245	6,823	20,083	35,582
3/86	-29,640	-14,608	-13,973	-8,735	41,533	64,948	60,984	42,496
4/86	-28,505	-17,702	-17,026	-11,762	43,024	62,252	58,305	39,725
5/86	-33,364	-26,777	-26,132	-20,804	38,522	53,577	49,576	30,941
6/86	-33,247	-29,927	-29,235	-23,892	38,998	50,828	46,852	28,112
7/86	-15,294	-15,082	-14,383	-9,025	57,313	66,077	62,084	43,239
8/86	-18,353	-24,174	-20,054	-18,053	54,617	57,391	56,795	34,472
9/86	-51,748	-63,583	-56,108	-57,447	21,586	18,390	21,125	-4,659
10/86	-48,991	-66,911	-55,985	-60,717	24,711	15,472	21,635	-7,666
11/86	-45,952	-70,213	-55,747	-64,304	28,118	12,581	22,260	-10,687
12/86	-25,345	-53,163	-36,946	8,300	47,519	28,371	39,812	40,045
.
8/88	67,731	63,518	68,366	66,342	146,413	151,616	150,849	126,509
9/88	34,308	24,517	31,137	27,319	113,383	113,056	114,033	87,824
10/88	37,434	21,597	30,123	24,408	116,905	110,579	113,433	85,250
11/88	40,733	18,716	29,208	21,499	120,602	108,142	112,934	82,682
12/88	39,514	7,258	20,256	37,845	117,706	95,148	102,544	81,348
1/89	83,847	76,335	78,919	100,870	162,039	164,225	161,206	185,669

^a MVHR and SDP refer to minimum variance hedging ratio and stochastic dynamic programming, respectively.

Table 3. Performance Comparison of Four Different Marketing Approaches over the August 1985 to December 1988 Period

(August 1985) Initial before- Tax Income	Accumulated after-Tax Profits at Dec. 1988	Mean of after-Tax Income	Standard Deviation of after-Tax Income	Before-Tax Termination Value in January 1989	Cash Position at Jan. 1989	Marginal Tax Rate for Equality with SDP
				(\$)		
All Cash						
-20,000	39,513	9,878	31,566	44,333	83,847	91
0.00	62,475	15,619	23,297	44,333	106,809	80
20,000	85,384	21,346	15,607	44,333	129,718	32
40,000	103,325	25,831	10,326	44,333	147,659	40
60,000	117,706	29,426	7,901	44,333	162,039	39
All Futures						
-20,000	7,258	1,814	39,630	69,077	76,335	NA*
0.00	30,261	7,565	31,292	69,077	99,338	NA
20,000	53,257	13,314	23,088	69,077	122,334	NA
40,000	76,246	19,061	15,233	69,077	145,323	74
60,000	95,147	23,787	9,657	69,077	164,225	61
Minimum Variance Hedging Ratio						
-20,000	20,256	5,064	33,915	58,663	78,919	503
0.00	43,054	10,764	25,495	58,663	101,717	460
20,000	65,852	16,462	17,204	58,663	124,514	256
40,000	87,437	21,859	9,808	58,663	146,100	55
60,000	102,543	25,636	5,917	58,663	161,206	54
Stochastic Dynamic Programming						
-20,000	37,844	9,461	11,258	63,026	100,870	
0.00	58,718	14,680	4,341	63,026	121,744	
20,000	72,654	18,163	3,516	63,026	135,679	
40,000	67,023	16,756	3,837	104,321	171,344	
60,000	81,348	20,337	4,563	104,321	185,669	

Note: Accumulated after-tax profits of the firm through December 1988, the mean and standard deviation of after-tax profits realized, the before-tax termination money received in January 1989 from closing out any outstanding futures position and selling all grain in storage on the cash market, the cash position of the firm at January 1989, and the marginal tax rate that would be required on all the excess before-tax termination income received in January from the stochastic dynamic programming (SDP) strategy over other strategies so that equal after-tax terminal wealth values would exist in January 1989.

* NA signifies not applicable because the SDP strategy has a lower before-tax termination value in January than this case.

compared, and this tax rate is still relatively high. Furthermore, the standard deviation of after-tax income is less for the SDP approach than all other marketing strategies considered. This result is undoubtedly due in part from the SDP framework incorporating the progressive income tax structure into grain marketing decisions. It also demonstrates that the variability of after-tax income is not increased from utilizing an SDP framework of grain-marketing strategies, even though risk aversion is not explicitly incorporated into the model. Thus, the SDP strategy outperformed all of the other strategies in accumulating wealth and minimizing risk.

Concluding Comments

Results of this study suggest that cash grain sales are most advantageous for grain producers when

basis levels are low and the before-tax income level of the firm is relatively low. Because of discounting and the progressive marginal tax structure, income tax considerations are most pronounced toward the end of the tax year. Therefore, at low cash price and low before-tax income levels, cash grain sales are relatively higher at the end of the tax year than at the beginning of the tax year. If the before-tax income level of the firm is relatively high, the producer can still take advantage of favorable cash price levels without increasing the current income tax liabilities of the firm by hedging grain on a futures contract that will mature in the following tax year.

Increased storage levels are more conducive to larger cash grain sales and increased hedging activity. However, cash grain sales tend to increase the most at lower before-tax income levels, and futures transactions tend to increase the

most at higher before-tax income levels. Increasing the short futures position of the firm generally results in decreased cash grain sales and futures transactions because of the legitimate hedging constraint (i.e., short futures position of the firm cannot exceed the level of grain storage plus anticipated production before maturity or forty thousand bushels).

Changes in optimal grain-marketing decisions from the 1986 Tax Reform Act would undoubtedly be to store less grain because of lower marginal income tax rates, as concluded in Tronstad and Taylor (1989b). However, the reduction in storage from tax reform would be less for the model utilized in this analysis which also considers hedging. This reduction is especially true for basis levels above approximately \$.35 per bushel, or whenever a combination of hedging and storage is indicated to be a favorable tool for reducing income tax liabilities.

Depending on the initial before-tax income level and marketing strategy, a stochastic dynamic programming (SDP) framework accumulated from \$5,961 to \$26,021 more wealth than all other grain-marketing strategies considered for the August 1985 to January 1989 period. Also, profits generated from the SDP approach would have been much larger if the 1988 drought had been less severe; the drought caused March futures contracts sold in May and June to be bought back at a much higher price. Furthermore, all other marketing strategies yielded a standard deviation of after-tax income from 30% to 621% greater than that from the SDP framework. Thus, the SDP approach to grain-marketing decisions outperformed more traditional grain-marketing approaches which do not consider the dynamic and stochastic elements associated with income taxes.

[Received July 1989; final revision received April 1990.]

References

- Adler, M., and J. Detemple. "Hedging with Futures in an Intertemporal Portfolio Context." *J. Futures Mkts.* 8(1988):249-69.
- Bellman, R. *Dynamic Programming*. Princeton NJ: Princeton University Press, 1957.
- Berck, P. "Portfolio Theory and the Demand for Futures: The Case of California Cotton." *Amer. J. Agr. Econ.* 63(1981):466-74.
- Bond, G. E., and S. R. Thompson. "Risk Aversion and the Recommended Hedging Ratio." *Amer. J. Agr. Econ.* 67(1985):870-72.
- Burt, O. R. "Dynamic Programming: Has Its Day Arrived?" *West. J. Agr. Econ.* 7(1982):381-94.
- Burt, O. R., and C. R. Taylor. "Reduction of State Variable Dimensions in Stochastic Dynamic Optimization Models Which Use Times-Series Data." *West. J. Agr. Econ.*, in press.
- Heifner, R. G. "Optimal Hedging Levels and Hedging Effectiveness in Cattle Feeding." *Agr. Econ. Res.* 24(1972): 25-36.
- Howard, R. A. *Dynamic Programming and Markov Processes*. New York: John Wiley & Sons for Massachusetts Institute of Technology, 1960.
- Johnson, L. L. "The Theory of Hedging and Speculation in Commodity Futures." *Rev. Econ. Stud.* 27(1960):139-51.
- Kahl, K. H. "Determination of the Recommended Hedging Ratio." *Amer. J. Agr. Econ.* 65(1983):603-5.
- Karp, L. "Methods for Selecting the Optimal Dynamic Hedge When Production Is Stochastic." *Amer. J. Agr. Econ.* 69(1987):647-57.
- Peck, A. E. "Hedging and Income Stability: Concepts, Implications, and an Example." *Amer. J. Agr. Econ.* 57(1975):410-19.
- Rolfo, J. "Optimal Hedging Under Price and Quantity Uncertainty: The Case of a Cocoa Producer." *J. Polit. Econ.* 88(1980):100-16.
- Schwarz, G. "Estimating the Dimension of a Model." *Ann. Statist.* 6(1978):461-64.
- Simon, H. "Dynamic Programming Under Uncertainty with a Quadratic Criterion Function." *Econometrica* 24(1956):74-81.
- Tronstad, R., and C. R. Taylor. *A Post-Sample Comparison of Stochastic Dynamic Programming Grain Marketing Decisions to Other Cash and Futures Trading Strategies*. Dep. Agr. Econ. Agr. Res. Rep. No. 40, University of Arizona, 1989a.
- . "Effects of the 1986 Tax Reform Act on Grain Marketing Decisions: A Case Study of Winter Wheat Producers." *N. Cent. J. Agr. Econ.* 11(1989b):309-20.
- Theil, H. "A Note on Certainty Equivalence in Dynamic Planning." *Econometrica* 25(1957):346-49.
- Ward, R. W., and L. B. Fletcher. "From Hedging to Pure Speculation: A Micro Model of Optimal Futures and Cash Market Positions." *Amer. J. Agr. Econ.* 53(1971): 71-78.
- Witt, H. J., T. C. Schroeder, and M. L. Hayenga. "Comparison of Analytical Approaches for Estimating Hedge Ratios for Agricultural Commodities." *J. Futures Mkts.* 7(1987):135-46.
- Working, H. "The Theory of Price of Storage." *Amer. Econ. Rev.* 39(1949):1254-62.

Price Changes, Supply Elasticities, Industry Organization, and Dairy Output Distribution

Adesoji O. Adelaja

Long-run price elasticities of milk supply are decomposed into price elasticities of yield, herd size, and farm population. For northeastern U.S. dairy farms, these elasticities are shown to vary with farm size because of size-related differences in capital intensity, specialization, yield and herd size variability, and rates of entry/exit into dairy. In the short run, price responsiveness is shown to decrease with farm size. In the long run, large farms are shown to be most price responsive. Results suggest that lower milk prices could make revenue distribution less equal in both the short and long run by altering output and population distributions.

Key words: dairy revenue distribution, herd size, population, supply elasticity.

In the postwar years, the federal milk price support program played an important role in U.S. milk pricing and output regulation. The program stimulated much of the increases in real milk prices, total dairy output, and total dairy herd size from the mid-1960s to the early 1980s (Gardner). The program was also partly responsible for the declines in real milk prices, total milk output and total dairy herd size in the U.S. since the 1983 Dairy and Tobacco Adjustment Act (Klemme and Chavas). Since the passage of the 1983 act, agricultural economists and policy makers have more actively debated the impacts of price support policies and declining milk prices on the dairy industry (Gardner).

To examine the impacts of milk price changes, Chavas and Klemme, Klemme and Chavas, and Kinnucan and Forker recently revisited earlier studies by Halvorsen; Chen, Courtney, and Schmitz; Wipf and Houck; and Levins on milk supply response. The more recent, as well as the older, studies focused only on aggregate short-and/or long-run price elasticities of supply but ignored the possibility that supply response could

vary with farm size. The only disaggregate study of responsiveness to price changes in the dairy industry focused on input demand adjustments of dairy farms (Hoque, Adelaja, and Ganguly).¹

A correlation between price responsiveness and farm size might exist for several reasons. First, because capital intensity and specialization are likely to increase with farm size, the ability to vary cost and therefore output in the short run (in response to a price change) should decrease with firm size (O_i). Second, because the financial ability of a farm and thus its ability to vary fixed inputs (e.g., herd size) beyond the short run (in response to a price change) are likely to be directly correlated with its size, long-run supply responses of dairy farms may not be homogenous. Third, by encouraging firm entry (exit) into (from) dairy (Baumol, Panzar, and Willig 1982; Baumol), a price change can affect firm population. The rates at which such entry and exit occur and the rates at which incumbents change their sizes (crossover entry/exit) may vary with farm size grouping because sunk costs, which are barriers to entry/exit, may vary with farm size (Dunne, Roberts, and Samuelson). Because the long-run supply response encompasses the effects of prices on farm population

Adesoji O. Adelaja is an assistant professor, Department of Agricultural Economics and Marketing, Rutgers University.

New Jersey Agricultural Experiment Station Publication No. D-02133-2-90. This research was supported by state funds and federal funds appropriated under the Hatch Act.

Review comments of associate editor John Braden and two anonymous *Journal* referees, the suggestions of Jim Seagraves, and Susan Howard's efficient word-processing assistance are greatly appreciated.

¹ Hoque, Adelaja, and Ganguly showed that input demand functions of dairy farms become more inelastic as farm size increases. Silberberg offered a theoretical explanation of the inverse relationship between input demand elasticities and firm size in competitive industries.

and farm size (Kislev and Peterson), it could also vary with herd size. Thus, aggregate short- and long-run supply elasticities may be misleading if they do not reveal farm-level supply responses.

By focusing on aggregate measures of short- and long-run milk supply elasticities, agricultural economists may have ignored several important effects of milk price changes. One effect involves the impacts of milk prices on the distribution of farm population by size. Knowledge of this is important in understanding the dynamics of dairy farm population, why farms enter and leave the dairy industry, and why some farm operators switch easily from farm to off-farm income. It could also be useful in planning government programs and services whose distribution is based on farm size. A second effect involves the differential impacts of milk prices on herd sizes and the relative differences among farms in potential herd expansion. Knowledge of these effects is useful in understanding the process of farm consolidation on dairy farms. Third are the differential impacts of milk prices on dairy output or revenue, which have equity implications for price related policies. Fourth are the differential impacts of milk prices on yield, which have implications for observed yield differences among farms. An understanding of all these aspects of price changes can enrich policy deliberations by enabling policy makers to consider both the aggregate and distributional effects of price changes and support policies.

The objective of this paper is to account for all the effects of dairy price changes discussed above. Specifically, the paper (a) conceptualizes the relationships between farm size and short-run/long-run output responsiveness, (b) presents a theoretical model for decomposing long-run supply elasticities into their components (the elasticities of yield, herd size, and dairy farm population), (c) presents a theoretical model for evaluating the output and revenue redistribution effects of price changes, (d) introduces the issue of equity into the price support debate, and (e) applies the models to panel data on a sample of northeastern U.S. dairy farms.

Supply Elasticity and Firm Size

This section conceptualizes the relationships between long-run elasticities of supply, short-run elasticities of supply, elasticities of yield, elasticities of herd size, and elasticities of population. It also conceptualizes how these elasticities

vary with farm size. The analysis draws on recent work on the effects of dairy price changes (Chavas and Klemme), on output flexibility theory (Mills and Schumann), and on contestability theory (Baumol).

If dairy farms cannot change the scales of their operations in the short run and if farm population cannot change in the short run, the only short-run outcome of a one-time price increase will be increased cow yield or output per cow (Chavas and Klemme).² This implies that the aggregate short-run price elasticity of supply (η) is equal to the aggregate short-run price elasticity of yield (γ). Chavas and Klemme showed that virtually all output adjustments (97%) made beyond the short run are accomplished through changes in total industry herd size (H). Thus, the short- and long-run price elasticities of yield are virtually equal, and output adjustments made beyond the short run are accomplished only through adjustments in the herd size of the average farm (\bar{H} or average herd size) and in the number of farms (N). This is because $H = (\bar{H})(N)$.

Baumol's theory of contestable markets and Kislev and Peterson's work on effects of price changes support the argument that output adjustments made beyond the short run involve increased farm population. That is, price increases (decreases) encourage entry and/or herd size expansion (exit and/or herd size contraction) if the market is competitive (Baumol, Kislev and Peterson). If one assumes that herd size proxies all fixed inputs, the difference between the measured value of η and the measured value of aggregate long-run supply elasticity (η^*) will be the aggregate price elasticities of average herd size (λ) and of farm population (β). That is, $\eta^* = \lambda + \beta + \gamma = \lambda + \beta + \eta$.³ The entire short-run adjustment in output will be captured in η , while the entire "beyond the short-run" output adjustment will be captured in $\lambda + \beta$.

Ladd and Martin cite the following reasons why $\eta^* \geq \eta$ or why $\lambda + \beta \geq 0$: (a) the fixity of some inputs in the short run, (b) imperfect knowledge or uncertainty, and (c) habit. First, because not all inputs are variable in the short

² Chavas and Klemme demonstrated this by showing that 100% of all short-run output adjustments of dairy farmers to a price change are accomplished through changes in yield.

³ Output (Q) = output per cow (\bar{Y}) times number of cows per farm (\bar{H}) times number of farms (N). Thus, $\partial \ln Q / \partial \ln P = \partial \ln \bar{Y} / \partial \ln P + \partial \ln \bar{H} / \partial \ln P + \partial \ln N / \partial \ln P$, where P is price and \ln is the natural log operator. Therefore, $\eta^* = \eta + \lambda + \beta$ and $\eta^* - \eta = \lambda + \beta$ if it is assumed that herd size and other fixed inputs are used in fixed proportion.

run (input fixity), the ability to vary output in the short run is more limited (limited to γ). Second, because producers may not be immediately aware of a price change (imperfect knowledge), output responsiveness is lower in the short run. Besides, producers may view short-run price changes as transient and therefore avoid costly investments or disinvestments such as herd size changes until it is confirmed that a price change is permanent. Third, producers may be so accustomed to current output and input levels (habit) that it takes some time to adjust output when price changes.

Assume that farms can be grouped by herd size and denote the disaggregate counterparts of the elasticities discussed above as η_i^* , η_i , γ_i , β_i and λ_i . The subscript i implies elasticities for farms in the i th size group. It will be shown below that each aggregate measure (η^* , η , λ , γ , or β) is the weighted average of its disaggregate counterpart (η_i^* , η_i , λ_i , γ_i , or β_i), whether or not the disaggregate one varies with farm size grouping. Note also that $\eta_i^* \geq \eta_i$ and that $\lambda_i + \beta_i \geq 0$, for each i (Chavas and Klemme, Ladd and Martin). The rest of this section conceptualizes how η_i^* , η_i , λ_i , γ_i , and β_i vary with herd size grouping.

Short-Run Supply Elasticity and Firm Size

In a competitive industry such as dairy, the short-run supply function for a firm is the portion of the marginal cost function that lies above the average variable cost function. The steeper the marginal cost function, the more inelastic the supply function and the more curved the average cost function (Stigler). Smaller firms will have flatter average cost functions and more elastic short-run supply functions, while larger firms have smaller η_i because they are less diversified and use more capital-intensive and specialized technologies (Mills and Schumann).^{4,5}

⁴ The reasons for this are as follows (Oi): Labor is relatively expensive, while capital is relatively inexpensive for large producers. The shadow value of managerial time is high for large firms because their managers have superior managerial ability and because their complex machinery-intensive technologies require specialized, highly paid, and highly skilled labor. Capital-intensive technologies have high adjustment costs, are less flexible, and are more efficient at high levels of output. Small firms choose labor-intensive technologies because wage rates they pay to the unskilled workers they employ are low.

⁵ Because some large dairy farms are diversified, an alternative hypothesis is that small farms are less output flexible. Output flexibility arising from diversification on large farms, however, is likely to be overshadowed by the inflexibility arising from low-capital intensity given that relatively few large farms are diversified (Tremblay).

These conditions hold for U.S. dairy, where large producers tend to utilize more capital-intensive technologies, more specialized technologies, and better skilled labor relative to smaller farmers. Besides, the evidence that increased capital/labor ratio, farm consolidation, and herd expansion have occurred simultaneously on dairy and other farms in the United States (Antle, Binswanger, Kislev and Peterson, Hoque and Adelaja) suggests that capital-intensive and inflexible technologies are synonymous with large-scale farming. Large dairy farms may be less capable of varying inputs, and thus output, in the short-run because more of their costs are fixed. The tendency of small dairy farms to rely more on off-farm income and to switch to off-farm employment when price falls (Adelaja and Rose; Salant, Smale, and Saupe) also suggests that small farms have larger η_i .

Long-Run Supply Elasticity and Firm Size

Both λ_i and β_i capture delayed (beyond the short run) reaction because of input fixity, uncertainty, imperfect knowledge, and habit (Ladd and Martin). Differences in λ_i and β_i among farms can be explained in part by differences in farmers' abilities to implement their reactions (degrees of fixity of assets, inputs, or technology). For example, the ability to vary herd size should be related to financial considerations. Larger and more capital-intensive farms, with greater borrowing capacity, should be more capable of varying herd size so that λ_i increases with herd size. If β_i is constant as size increases, the proportion of the total adjustments to a price change occurring in the short run would be greater for small farms than for large farms. Also, while η_i^* s would be more similar for small and large farms than η_i s, the adjustment period would be shorter for small farms.

To examine the relationship between herd size and β_i , suppose the market is perfectly competitive and thus perfectly contestable (Baumol). The industry will be vulnerable to hit-and-run entry because firms can enter when price increases and exit possibly even before price decreases.⁶ The probabilities as well as rates of entry (exit) into (from) farming at various levels

⁶ A purely competitive market is purely contestable (Baumol). Free entry and costless exit imply no cost discrimination against entrants or exitants beyond normal user costs and depreciation. Potential entrants are at no risk of any retaliatory price adjustment action by incumbent firms, and economic profit will be zero for all firms at long-run competitive equilibrium.

of size should be similar (β_i is constant but > 0) if there are no economies-of-scale advantages to large-scale production.⁷

In real life, sunk costs usually represent barriers to free entry and costless exit (Baumol and Willig; Baumol, Panzar, and Willig 1983; Weitzman). Sunk costs, which are a proportion of fixed costs (Baumol, Panzar, and Willig 1983), should increase with size. This relationship and the direct correlation between the degree of risk and firm size (Hall and Weiss, Mills and Schumann) suggest that β_i decreases with firm size if returns to scale are constant (CRS).

When long-run average cost is declining, the relationship between β_i and herd size is less clear. Increasing returns to scale, by itself, implies that β_i increases with firm size. For example, Kesides showed that entry increases with the degree of profitability of incumbents. Panzar and Willig's analysis also suggests this relationship. If profitability increases with firm size (Hall and Weiss) as a result of economies of scale, entry or exit encouraged by a given price change will increase with firm size if there are no sunk costs.

The evidence on economies of scale in dairy farming (particularly Northeastern dairy) suggests almost constant returns to scale (Hoque and Adelaja; Matulich; Hoque, Adelaja, and Ganguly) and an inverse relationship between the elasticity of population (β_i) and farm size. Small farms' greater reliance on off-farm income and their greater flexibility in switching from farm to off-farm income further suggests that these farms may be more capable of hit-and-run entry.⁸ In conclusion, the expectation for the dairy industry is that short-run supply elasticity (η_i) and population elasticity (β_i) decrease with farm size while the elasticity of average herd size (λ_i) increases with farm size, but that $\eta_i^* = \eta_i + \lambda_i + \beta_i$. The long-run impact of price changes on output is decomposed into the short-run impact (η_i), the elasticity of average herd size (λ_i), and the elasticity of population (β_i) in the next section to allow the examination of the impact of price changes on output distribution.

Price Changes and Output Distribution

Assume that dairy farms can be grouped into n groups based on herd size (e.g., 0–39, 40–79,

and 80 and above milk cows). Further, assume that milk cows and all other fixed inputs are used in fixed proportions so that herd size proxies all fixed inputs or enterprise scale. Denote total output of farms in the i th group as Q_i and total output of all farms as Q . Furthermore, assume that each of the i th groups has m_i farms such that the total number of farms in the industry (N) is $\sum_{i=1}^n m_i$ and the population share of farms in the i th group (b_i) is m_i/N . Thus, $Q_i = m_i \bar{Q}_i$ and $Q = \sum_{i=1}^n Q_i = \sum_{i=1}^n m_i \bar{Q}_i$, where \bar{Q}_i is the output of the average farm in the i th group.

Denote the yield of the average farm in the i th group as \bar{Y}_i , the total number of milk cows on farms in the i th group as H_i , and the herd size of the average farm in the i th group as \bar{H}_i , such that $H_i = m_i \bar{H}_i$. Total herd size in the industry (H) is $\sum_{i=1}^n H_i = \sum_{i=1}^n m_i \bar{H}_i$ and the i th group's share of $H(a_i)$ is $H_i/H = m_i \bar{H}_i/H$. Note that $\bar{Q}_i = \bar{H}_i \bar{Y}_i$, that $Q_i = m_i \bar{H}_i \bar{Y}_i$, and that average industry herd size (\bar{H}) is H/N . Total industry output is

$$(1) \quad Q = \sum_{i=1}^n Q_i = \sum_{i=1}^n m_i \bar{H}_i \bar{Y}_i.$$

The share of total output contributed by farms in the i th group (S_i) is

$$(2) \quad S_i = \frac{Q_i}{Q} = \frac{m_i \bar{H}_i \bar{Y}_i}{Q},$$

where $\sum_{i=1}^n S_i = 1$. Assuming further that all producers are price takers who receive the same price (P), the total aggregate supply elasticity (Π) is⁹

$$(3) \quad \Pi = \frac{\partial \ln Q}{\partial \ln P} = \sum_{i=1}^n \frac{m_i \bar{H}_i \bar{Y}_i}{Q} \cdot \left(\frac{\partial \ln \bar{Y}_i}{\partial \ln P} + \frac{\partial \ln m_i}{\partial \ln P} + \frac{\partial \ln \bar{H}_i}{\partial \ln P} \right) \\ = \sum_{i=1}^n S_i (\gamma_i + \lambda_i + \beta_i) \\ = \sum_{i=1}^n S_i \gamma_i + \sum_{i=1}^n S_i \lambda_i + \sum_{i=1}^n S_i \beta_i.$$

In (3), Π includes the farm population (entry and exit from outside the industry, and cross-over type of entry and exit), herd expansion (increase

⁷ Dunne, Roberts, and Samuelson provided empirical evidence that entry and exit rates are directly correlated and that both are higher for small producers *vis-à-vis* large ones.

⁸ Entry and exit could be so frictionless for small producers that they contribute to short-run output flexibility (Baumol).

⁹ Since $Q = \sum_{i=1}^n Q_i$, $dQ = \sum_{i=1}^n dQ_i$, $dQ/Q = d \ln Q = \sum_{i=1}^n (Q_i/Q) (dQ_i/Q_i) = \sum_{i=1}^n S_i d \ln Q_i$. Therefore, $\partial \ln Q / \partial \ln P = \sum_{i=1}^n S_i \partial \ln Q_i / \partial \ln P$. Since $Q_i = m_i \bar{H}_i \bar{Y}_i$, $\partial \ln Q_i / \partial \ln P = \partial \ln m_i / \partial \ln P + \partial \ln \bar{H}_i / \partial \ln P + \partial \ln \bar{Y}_i / \partial \ln P$. Therefore, $\partial \ln Q / \partial \ln P = \sum_{i=1}^n S_i [(\partial \ln m_i / \partial \ln P) + (\partial \ln \bar{H}_i / \partial \ln P) + (\partial \ln \bar{Y}_i / \partial \ln P)] = \sum_{i=1}^n S_i (\beta_i + \lambda_i + \gamma_i)$.

in herd size not sufficient to create crossover), and yield (increase in output at constant levels of herd size) effects of price changes.¹⁰

Because $N = \sum_{i=1}^n m_i$, the aggregate price elasticity of population (β), which measures the net change in total farm population from a price change, is

$$(4) \quad \beta = \frac{\partial \ln N}{\partial \ln P} = \sum_{i=1}^n \frac{m_i}{N} \frac{\partial \ln m_i}{\partial \ln P} \\ = \sum_{i=1}^n b_i \beta_i \neq \sum_{i=1}^n S_i \beta_i,$$

where $\sum_{i=1}^n b_i = 1$. Note that $b_i \neq S_i$. Also, because $H = \sum_{i=1}^n m_i \bar{H}_i$ and $\bar{H} = \sum_{i=1}^n (m_i \bar{H}_i / N)$, the aggregate price elasticity of average herd size (λ), which measures the net change in average industry herd size due to a price change, is

$$(5) \quad \lambda = \frac{\partial \ln \bar{H}}{\partial \ln P} = \sum_{i=1}^n \frac{\bar{H}_i}{\bar{H}} \frac{\partial \ln \bar{H}_i}{\partial \ln P} \\ = \sum_{i=1}^n \frac{H_i}{H} \frac{\partial \ln \bar{H}_i}{\partial \ln P} = \sum_{i=1}^n a_i \lambda_i \neq \sum_{i=1}^n S_i \lambda_i,$$

where $\sum_{i=1}^n a_i = 1$ and $a_i \neq S_i \neq b_i$. Also, define the industry's average yield (\bar{Y}) as $\sum_{i=1}^n S_i \bar{Y}_i$ and obtain the price elasticity of aggregate industry yield (γ), which measures the net increase or decrease in yield (via increased output at constant fixed inputs) as a result of a price change, as

$$(6) \quad \gamma = \frac{\partial \ln \bar{Y}}{\partial \ln P} = \sum_{i=1}^n S_i \frac{\partial \ln \bar{Y}_i}{\partial \ln P} = \sum_{i=1}^n S_i \gamma_i.$$

Now, consider what γ_i , λ_i , and β_i measure when estimated from grouped data on a sample of farms. The value of γ_i would measure the increase in yield as a result of a price increase for a representative farm in the i th group when herd size is held constant. Increases in yield when herd size is held constant does not imply crossover. On the other hand, λ_i measures the increase in the herd size for the same farm when crossover does not happen. Such an increase accrues beyond the short run (Chavas and Klemme). Because some farms are located close to their boundaries, a change in herd size can move them into the next herd size category and can affect the population of the next category (crossover entry and exit). β_i is a convolution of the effects of price changes on population via crossover en-

try and exit (by incumbents) and entry and exit in and out of the industry.

Because the group with the largest farms is unbounded from above (e.g., 80 cows to infinity), the value of β_i obtained econometrically from grouped data for this group may be high (even higher than β_i for other groups) because farms can only enter when price increases. This, however, would not necessarily imply an upward bias in the value of β_i . The high β_i value for this group would simply reflect the absence of an upper bound.

Considering that $S_i = m_i \bar{H}_i \bar{Y}_i / Q$, the price elasticity of output share (ϵ_i) is

$$(7) \quad \epsilon_i = \frac{\partial \ln S_i}{\partial \ln P} = \frac{\partial \ln m_i}{\partial \ln P} + \frac{\partial \ln \bar{H}_i}{\partial \ln P} \\ + \frac{\partial \ln \bar{Y}_i}{\partial \ln P} - \frac{\partial \ln Q}{\partial \ln P} = \beta_i + \lambda_i + \gamma_i - \Pi \\ = \beta_i + \lambda_i + \gamma_i - \sum_{i=1}^n S_i (\beta_i + \lambda_i + \gamma_i) \\ = \gamma_i - \gamma + \sum_{i=1}^n (1 - S_i) (\beta_i + \lambda_i).$$

Price Changes and Output Distribution in the Short Run

In the short run, one can assume that $\beta_i = 0$ and $\lambda_i = 0$ (Chavas and Klemme). In this case, equation (3) converges into $\Pi = \sum_{i=1}^n S_i \gamma_i$. Because $\gamma_i = \eta_i$ and $\gamma = \eta$ (Chavas and Klemme), the following is obvious:

$$(8) \quad \eta = \sum_{i=1}^n S_i \eta_i.$$

Equation (8) shows that the aggregate short-run supply response correctly predicts the farm-level short-run supply elasticity ($\eta = \eta_i \sum_{i=1}^n S_i = \eta_i$) only if all farms have similar η_i .

The short-run price elasticity of output share (Y_i) is obtained from equation (7) as

$$(9) \quad Y_i = \gamma_i - \gamma = \eta_i - \eta.$$

Equation (9) shows that if η_i is greater (less) than η , a price increase raises (lowers) the output share for the i th group. If, however, $\eta_i = \eta$, $Y_i = 0$ and price changes will not affect short-run output distribution. The necessary and sufficient condition for price changes to affect short-run output distribution is that η_i varies with farm size. Note that the magnitude of Y_i increases as the differences in η_i among farms increase. The

¹⁰ Crossover entry is defined as entry into a group as a result of exit from another group.

implications of equations (8) and (9) are the same for revenue distribution if the competitive pricing assumption is maintained.¹¹

Price Changes and Output Distribution in the Long Run

In the long run, β_i and λ_i are expected to be positive, λ_i is expected to increase with farm size and β_i is expected to decrease with farm size but only up to a point. The following long-run relationship between aggregate long-run supply elasticity and farm-level supply elasticity is obvious from equation (3):

$$(10) \quad \eta^* = \sum_{i=1}^n S_i \eta_i + \sum_{i=1}^n S_i \lambda_i + \sum_{i=1}^n S_i \beta_i \\ = \sum_{i=1}^n S_i (\eta_i + \lambda_i + \beta_i) = \sum_{i=1}^n S_i (\eta_i).$$

From equation (10), if there are differences in η_i^* , η^* will not correctly predict farm level supply response. The differences in η_i^* may come from differences in η_i , λ_i and γ_i . Even when the η_i s are similar, η^* will still not correctly predict farm level supply response as long as λ_i and β_i vary with farm size.

The long-run price elasticity of output share (Y_i^*) is obtained from equation (7) as follows:

$$(11) \quad Y_i^* = \gamma_i + \lambda_i + \beta_i \\ - \sum_{i=1}^n S_i (\gamma_i + \lambda_i + \beta_i) = \eta_i^* - \eta^* \\ = \eta_i - \eta + \sum_{i=1}^n (1 - S_i) (\lambda_i + \beta_i) \\ = Y_i + \sum_{i=1}^n (1 - S_i) (\lambda_i + \beta_i).$$

From (11), if the i th farm group's η_i^* is greater (less) than η^* , a price increase raises (lowers) the output share for the i th group in the long run. The requirement for price changes not to affect long-run output distribution is that all farms have similar η_i^* . Even if η_i s are similar ($Y_i = 0$), redistribution of output still occurs in the long

run if there are differences in λ_i and β_i . Note that if η_i^* are more similar than η_i , there may be an over redistribution of output in the short run which could be corrected for in the long run.

In evaluating the revenue redistribution effects of price policies, values of Y_i and Y_i^* derived from estimates of η_i , η_i^* , λ_i , and β_i can be used to project revenue and output shares that result from a given price change. They can also be used to construct Lorenz-type curves for output or revenue distribution (by plotting the cumulative percentage of farms against the cumulative percentage of output or revenue before and after an anticipated price increase or decrease). Finally, they can be used to calculate measures of output or revenue inequality such as the gini coefficient (G).

Panel data are ideal for estimating η_i , η_i^* , λ_i , and γ_i , and thus Y_i and Y_i^* for farms of different sizes. In this study, data on average farms in several size groups collected over time are used. Each time series on the average farm in a group is used to estimate the supply function for the group the farm represents. The short- and long-run price elasticities of the product of the output of the average farm in a group and the population of the group were used to approximate the weighted short- and long-run supply elasticities for the group.¹² This method is valid if the distribution of farms in the sample used approximates the distribution of the total population.

Evidence From Northeastern U.S. Dairy

The decomposition model above was applied to data available through the electronic farm accounting (ELFAC) system. This system is sponsored by the Cooperative Extension Service for the northeastern states. The system provides annual data on average farms in several groups for a sample of northeastern U.S. dairy farms. In annual ELFAC reports (Tremblay) available for the 1971 through 1985 period, average production and cost data are reported for small farms (less than 40 milk cows), medium-sized farms (40 to 79 milk cows) and large farms (80 or more milk cows). Specific variables upon which data are reported include herd size, dairy revenue, nondairy revenue, variable production ex-

¹¹ To evaluate fully the equity implication of price changes, the impact of price changes on profit distribution would be needed. This impact would require knowledge of the structures of the underlying cost functions for farms in various size groups and would introduce much greater methodological and analytical complexity. The discussion in this paper is limited to revenue distribution but the extension of this methodology to assess profit distribution is left to future research.

¹² Since $\bar{Q}_i = Q_i/m_i$, $m_i \bar{Q}_i = a_i \sum_{j=1}^m Q_{ij}$ where Q_{ij} is the output of the j th farm in the i th group. Therefore, $\partial \ln (m_i \bar{Q}_i) / \partial \ln P = \sum_{j=1}^m (Q_{ij}/m_i \bar{Q}_i) (\partial \ln Q_{ij} / \partial \ln P) = \sum_{j=1}^m r_{ij} (\partial \ln Q_{ij} / \partial \ln P)$ where $r_{ij} = Q_{ij}/m_i \bar{Q}_i$. This shows that the price elasticity of the product of population and the output of the average farm in the i th group is the weighted average supply elasticity for farms in the i th group.

penses, price received, milk sale in pounds, number of family workers, value of land, value of buildings, value of livestock, and value of equipment.

For each group, the mean of the annual values of herd size of the average farm (\bar{H}_i) and the coefficient of variation ($CV_{\bar{H}_i}$) were examined to determine variability of herd sizes. In the small herd size group, mean \bar{H}_i was 33 (maximum = 35, minimum = 32) and $CV_{\bar{H}_i}$ was 2.48%. In the medium herd size group, mean \bar{H}_i was 57 (maximum = 58, minimum = 55) and $CV_{\bar{H}_i}$ was 1.55%. In the large herd size group, mean \bar{H}_i was 122 (maximum = 130, minimum = 116) and $CV_{\bar{H}_i}$ was 3.72%. Mean \bar{H}_i for all farms was 71 (maximum = 75, minimum = 65) while $CV_{\bar{H}_i}$ was 4.43%. The size of the average farm seemed relatively stable over time. The higher $CV_{\bar{H}_i}$ value for large farms probably reflects the fact that herd size is unbounded from above. Because of the nature of the data, changes in \bar{H}_i indicate only herd expansion/contraction (λ_i) that do not involve crossover. Herd size changes involving crossover are captured in β_i .

The Koyck distributed lag specification was chosen over the Nerlove (partial adjustment) and Almon for estimating short- and long-run supply elasticities. Both the Koyck and the Nerlove are similar except in the structures of the stochastic disturbance terms and the interpretations of the adjustment paths of output. The Nerlove presumes that the search for new optimality after a price increase involves periods of overproduction and underproduction relative to the new optimal output level. Periods of overproduction seem unlikely given the reasons λ_i and β_i are expected to be positive. Both the Almon and the Koyck presume that continuous underadjustment ensues whereby output levels during adjustment approach the new optimal level as a limit. However, while the Koyck assumes that the proportion of adjustments made in a given time declines over time, the Almon assumes that the proportion increases and then decreases. The Koyck specification was used because of the expected short-run stickiness of population and herd size and the limited degrees of freedom imposed by the data.

The following Koyck supply function was specified for each group of farms:

$$(12) \quad \ln QM_{it} = \alpha_{i0} + \alpha_{i1} \ln PM_{it} + \alpha_{i2} \ln PF_{it} + \alpha_{i3} \ln PB_{it} + \alpha_{i4} \ln QM_{it-1} + U_{it},$$

where QM_{it} is the product of mean milk output (output of average farm) and number of dairy

farms in the i th group in time period t , PM_{it} is the real average price received by farmers in the i th group for milk in time period t , PF_{it} is the real weighted average price paid by farmers in the i th group for feed and concentrates in time period t , PB_{it} is the real average price received by farmers in the i th group for beef (a substitute product) in time period t , QM_{it-1} is the lagged value of the dependent variable (QM_{it}), and U_{it} is a stochastic error term. The coefficients of the models are α_{i0} , α_{i1} , α_{i2} , α_{i3} , and α_{i4} .

The model assumes that the effects of a one-time increase in PM , PF , and PB are distributed over time and that the weights decline geometrically. The estimate of short-run supply elasticity (η_i) is α_{i1} , the estimate of supply elasticity after the T th period is $\alpha_{i1}(1 - \alpha_{i4}^T)/(1 - \alpha_{i4})$, the estimate of long-run supply elasticity (η_i^*) is $\alpha_{i1}/(1 - \alpha_{i4})$, the mean lag is $\alpha_{i4}/(1 - \alpha_{i4})$, and the median lag is $\ln 0.5/\ln \alpha_{i4}$. The specification imposes the same lag structure on the impact of PM_t , PF_t , and PB_t on QM_t for farms in the same herd size group. Lag structures, however, are free to vary over the herd size groups.

The QM_{it} variable for a group was measured as the annual mean of output in hundredweights times the number of farms in the group because the distribution of the ELFAC sample is very similar to that of the northeastern population. PM_{it} was measured as the deflated values of mean prices received by ELFAC farmers. PF_{it} and PB_{it} were measured as the deflated values of the indexes of nominal prices paid by farmers for feed and received by farmers for beef in the Northeast (*Agricultural Prices*). The deflator used for all prices was the deflator for gross national product (DGNP) obtained from *The Economic Report of the President*.

To explain the relationships between λ_i , β_i , η_i , and η_i^* , a specialization index (SI_i) and a capital intensity index (CI_i) were calculated from the ELFAC data set. Annual values of SI_i for the average farms in each group were calculated as the percentages of total revenue derived from dairy-related activities. From the annual values of SI_i , mean values of \bar{SI}_i (\overline{SI}_i) were calculated. The estimated values of \bar{SI}_i for small, medium, large, and all farms were, respectively, 92.52, 94.20, 93.80, and 94.00. Because the values of \bar{SI}_i were significantly different at the 5% level for small and medium and for small and large farms (but not for medium and large farms), it was concluded that specialization increases very slightly with herd size but only up to a point.

Total variable cost (TVC_i) and total fixed cost (TFC_i) for the average farms were calculated to

obtain annual values of CI_i . TVC_i was calculated as the sum of expenses on all variable inputs including hired labor. The components of fixed costs were assumed to be fixed investments, family labor, and property taxes. Cost of family labor was calculated as the number of family workers times the agricultural wage rate per hour in the Northeast (*Farm Labor*) times forty hours a week times fifty-two weeks a year. Thus, the productivity of family and hired labor were assumed equal. Following Christensen and Jorgenson, the opportunity cost of fixed investments was calculated as the product of interest rate paid by farmers (*Agricultural Statistics*) and total asset value. TFC_i was calculated as the sum of the opportunity cost of fixed investment, family labor cost, and property taxes paid.

Total production cost (TC_i) was measured as $TVC_i + TFC_i$. Annual values of CI_i were obtained as the percentages of total cost that are fixed. The means of these annual capital intensity values ($\overline{CI_i}$) obtained for small, medium, large, and all farms were, respectively, 0.172, 0.202, 0.223, and 0.188. All were significantly different at the 5% level, supporting Mills and Schumann's argument that the degree of capital intensity increases slightly with herd size. Output share (S_i) measures calculated from the EL-FAC data set for 1985 were, respectively, .0726, .4251, and .5023 for small, medium, and large dairy farms. These are similar to the shares reported by the Agricultural Statistics Board (*Milk Production*) for the entire U.S. dairy industry.

Econometric Considerations

Ordinarily, serial correlation makes ordinary least squares (OLS) estimates unbiased but inefficient. With the Koyck specification and the introduction of a lagged dependent variable, OLS estimates are biased, inconsistent and inefficient (Pindyck and Rubinfeld) because the lagged dependent variable and the error term will be correlated and the correlation will not disappear as sample size increases (Griliches).¹³ Generalized differencing methods are not appropriate for the Koyck model (Wold). Generalized differencing would lead to biased estimates while techniques for eliminating the bias will lead to inefficient parameter-estimates (Pindyck and Rubinfeld).

The instrumental variable and maximum like-

lihood methods have been recommended for estimating Koyck models (Johnston). The former is chosen over the latter whose estimators are not exact maximum likelihood estimators but maximum likelihood conditional on the first few observations. The instrument used to replace the lagged dependent variable is the predicted value of the dependent variable based on a regression of the dependent variable on current and lagged values of the independent variables (excluding the lagged dependent variable). The instrumental variable models were estimated via generalized least squares.¹⁴

Results

Parameter estimates of the supply functions for small, medium, large, and all farms are reported in table 1. All α_{i1} coefficients are significant at the 5% level. All α_{i2} coefficients are significant at the 5% level with the exception of the α_{i2} for small farms which is significant only at the 10% level. All α_{i3} coefficients are insignificant at both the 10% and 5% levels, suggesting the irrelevance of beef prices in the determination of milk supply. All the α_{i4} coefficients are significant at the 5% level. The signs of all significant coefficients are consistent with expectations.

For each model, Glegser's test for heteroscedasticity was performed by regressing the absolute values of the residuals against the dependent variable. The hypothesis that heteroscedasticity was present was rejected in each case. Shapiro-Wilkes tests for nonnormality of the error terms also suggested normal distributions of the residuals.

The estimated values of η_i in table 2 support the hypothesis that short-run supply elasticity decreases with farm size. The values are also consistent with the capital intensity and specialization indexes, suggesting that small farms may have greater ability to vary feeding rates, concentrate ratios, and the levels of labor utilization, and to vary yield and output in the short run. As expected, median and mean lags are directly correlated with farm size. Also, as shown in table 2, the proportions of total adjustments in the output to a one-time price change completed in one year were 77% on small farms, 70% on medium sized farms, and 30% on large farms.

Estimated values of η_i for small, medium, and

¹³ The Durbin-Watson d -statistic is inappropriate when a regression contains a lagged dependent variable. The Durbin L -statistic is recommended for large samples.

¹⁴ The Proc Autocor default procedure in SAS was used.

Table 1. Parameter Estimates of Supply Functions for ELFAC Dairy Farms, 1971–85

Dependent Variable = $\log QM_t$		Farm Size Groups ^a			
Independent Variable	Coefficient	Small	Medium	Large	All
Intercept	α_0	15.7078** (6.0830)	15.7243** (4.9906)	4.7620 (3.5106)	15.5419** (5.9349)
Log milk price (PM_t)	α_{11}	0.5204** (0.1480)	0.2665** (0.0890)	0.2311** (0.1010)	0.3646** (0.1048)
Log feed price (PF_t)	α_{12}	-1.2037** (0.5156)	-1.0635** (0.3163)	-0.6689* (0.3285)	-1.0492** (0.3891)
Log beef price (PB_t)	α_{13}	0.0178 (0.3771)	-0.0406 (0.2679)	-0.0284 (0.3177)	0.0231 (0.2916)
Log lagged quantity (QM_{t-1})	α_{14}	0.2331** (0.1135)	0.3012** (0.1347)	0.6954** (0.2312)	0.2896** (0.1247)
Regression R-square		0.3989	0.6774	0.6990	0.6723
Total R-square		0.5365	0.6800	0.7585	0.6942

^a Small farms = less than 40 milk cows; medium-sized farms = 40–79 milk cows; large farms = over 79 milk cows; and all farms = small, medium, and large farms combined. Double and single asterisk indicate significance at the 5% and 10% levels, respectively. Standard errors are parenthesized. Estimates generated by OLS. PM_t is real price of milk in period t ; PF_t and PB_t are real prices of feed and beef, respectively, in period t ; and QM_{t-1} is lagged value of QM_t . QM_t is the product of mean group output and group size in time t in hundredweights.

large sized farms are, respectively, 0.5204, 0.2665, and 0.2311. The aggregate short-run elasticity (η) of 0.3646, estimated from the data on all farms, exceeds η_i for medium and large farms but is less than η_i for small farms. The estimate of η is similar to Halvorsen's estimate of 0.30 for the United States, higher than Chavas and Klemme's estimate of 0.11 for the United States, higher than Chen, Courtney, and Schmitz's estimate of 0.16 for California, and lower than Levins' and Milligan's estimates of 0.91 and 0.63 for Mississippi. The results of this study are consistent with those of previous studies considering that dairy farms in California are

larger than those in the Northeast, which are in turn larger than those in Mississippi. However, it is somewhat contradictory that this study's estimates suggest more output flexibility than that of Chavas and Klemme, considering that dairies in the Northeast are generally smaller than in the rest of the United States.

Estimated values of η^* for small, medium, and large sized farms are, respectively, 0.6785, 0.3814, and 0.7585. The aggregate long-run elasticity of 0.5132 is lower than Levins', Milligan's, and Chavas and Klemme's estimates of 1.55, 1.57, and 6.69, respectively. The difference between the short-run and long-run supply

Table 2. Elasticities of Milk Supply for ELFAC Dairy Farms, 1971–85

	Farm Size Category			
	Small ^a	Medium	Large	All
Own-price elasticity:				
Instantaneous (η_i) ^b	0.5204	0.2665	0.2311	0.3646
After 1 year	0.6417	0.3468	0.3918	0.4702
After 2 years	0.6700	0.3708	0.5035	0.5008
After 3 years	0.6765	0.3782	0.5812	0.5096
After 4 years	0.6781	0.3804	0.6353	0.5122
After all adjustments (η^*) ^c	0.6785	0.3814	0.7585	0.5132
Lag characteristics:				
Mean lag	0.1780	0.2105	2.2830	0.4077
Median lag	0.4760	0.5776	1.9081	0.5593
Percent of adjustments completed instantaneously	77	70	30	71
Percent of adjustments completed in 1 year	95	91	52	91
Percent of adjustments completed in 2 years	98	97	66	98

^a Categorization of farms is same as in table 1.

^b Short-run elasticity.

^c Long-run elasticity.

Table 3. Summary of Results From Yield, Herd Size, and Population Equations

Estimate	Farm Size Category			
	Small	Medium	Large	All
Elasticity of yield (γ_i)				
Short-run	0.5018***	0.2504**	0.2337**	0.3551**
Long-run ^b	0.5018	0.2504	0.2337	0.3551
Difference				
Elasticity of herd size (λ_i)				
Short-run	0.0120**	0.0141**	0.0214**	0.0193**
Long-run ^b	0.0313	0.0519	0.2131	0.1255
Difference	0.0193	0.0378	0.1917	0.1162
Elasticity of population (β_i)				
Short-run	0.0561**	0.0254**	0.0423**	0.0417**
Long-run ^c	0.1820	0.0775	0.3352	0.0735
Difference	0.1259	0.0521	0.2929	0.0318
Implied value of η_i^*	0.7151	0.3798	0.7820	0.5541

¹ Double and single asterisks imply significance at the 5% and 10% levels, respectively.

² Short-run and long-run price elasticities of yield assumed to be equal because the coefficients of the lagged dependent variables were insignificant at the 5% and 10% levels.

³ Long-run elasticity calculated as short-run elasticity divided by $(1 - \text{coefficient of the lagged dependent variable})$.

elasticities for medium-sized farms ($\lambda_i + \beta_i = 0.1140$) suggests that these farms do not make significant short-run yield adjustments, and do not make significant long-run herd size or population adjustments. The sources of short- and long-run inflexibility on mid-sized dairy farms may need further investigation. Small farms are very price responsive in the short run but do not make much additional adjustment in the long run ($\lambda_i + \beta_i = 0.1581$). Large farms, however, which are not very price responsive in the short run are even more price responsive than the small farms in the long run ($\lambda_i + \beta_i = 0.5274$). Large farms' flexibility comes from herd size expansion and population change. For all farms, $\lambda_i + \beta_i = 0.1486$, which suggests relatively little herd size and population adjustments overall.

The ideal procedure for decomposing $\eta_i^* - \eta_i$ into its components (λ_i and β_i) is to estimate equation (12) jointly with three other equations similar to it, but where $\ln QM_{it}$ is replaced with $\ln(\bar{Y}_i)$, $\ln(\bar{H}_i)$, and $\ln(\bar{m}_i)$ and to constrain the sums of similar elasticities in the other equations to the elasticity in equation (12). However, joint estimation of a system of seemingly unrelated equations with constraints and lagged dependent variables is complex. Instead, the yield, average herd size, and population models were estimated individually and without constraints using the instrumental variable method used to estimate equation (12). A summary of the results is presented in table 3.

All short-run price elasticities of yield (γ_i) were similar to the short-run price elasticities of supply (η_i). Also, the coefficients of the lagged de-

pendent variable in the yield equations were all insignificant. These results confirm the hypothesis that virtually all short-run output adjustments are accomplished through yield changes ($\gamma_i \approx \eta_i$). The observed inverse relationship between γ_i and herd size suggests that cow production efficiency (yield) measures obtained at any given point in time are dependent on prices. They also suggest more variability in the yield of smaller farms, relative to larger ones.

From the herd size equation, the coefficient for price was 0.0120 for small farms, 0.0141 for medium-sized farms and 0.0214 for large farms, while the long-run price elasticities of herd size were 0.0313, 0.0519, and 0.2131, respectively. These values suggest increasing λ_i as herd size increases, that virtually no herd size adjustment is accomplished in the short run, and that price increases stimulate little herd size expansion even in the long run on small and medium-sized farms. The high value of λ_i for large farms probably reflects the fact that this group is unbounded from above. For all farms, the short- and long-run elasticities of herd size were, respectively, 0.0193 and 0.1255.

From the population equation, the coefficients for price were 0.0561 for small farms, 0.0254 for medium-sized farms, and 0.0423 for large farms, while the long-run price elasticities of population were 0.1820, 0.0775, and 0.3352, respectively.¹⁵ These results suggest that the

¹⁵ Tests for heteroscedasticity and nonnormality were also performed on the yield, herd size, and population equations. Error terms in all equations were also found to be homoscedastic and normal.

Table 4. Percentage Changes in Revenues and Revenue Shares from a 50% Decrease in Milk Price

Measure of	Measured as	Small	Medium	Large	All
Initial revenue shares in 1985 (%)	S_i	7.26	42.51	50.23	100.00
Supply elasticity (SR)	η_i	0.5204	0.2665	0.2311	0.2672 ^a
Supply elasticity (LR)	η_i^*	0.6785	0.3814	0.7585	0.5924 ^a
Elasticity of output or revenue share (SR) ^b	Y_i	0.2532	-0.0007	-0.0361	
Elasticity of output or revenue shares (LR) ^b	Y_i^*	0.0861	-0.2110	0.1661	
Short-run impact					
Change in output or revenues (%)		-26.02	-13.33	-11.56	-13.36
Change in output or revenue share (%)		-12.66	+0.35	+1.805	
Final output or revenue share (%)		6.34	42.65	51.14	
Long-run impact					
Change in output or revenue (%)		-33.93	-19.07	-37.93	-29.62
Change in output or revenue share (%)		-4.305	+10.55	-8.305	
Final output or revenue share (%)		6.94	47.00	46.06	

^a η and η^* are calculated as $\sum_{i=1}^n S_i \eta_i$ and $\sum_{i=1}^n S_i \eta_i^*$ and not as elasticities obtained from estimating supply function for all farms. This is because the estimated and calculated elasticities differ slightly.

^b SR and LR are notations for short run and long run, respectively.

population effect occurs primarily in the long-run and that β_i decreases with herd size, but only until the category with the largest farms is reached. As explained above, β_i for that group is expected to be high because the group's range is unbounded. This problem cannot be completely eliminated even when one breaks this extreme group into smaller groups (there will always be one group with an unbounded range). Fichtenbaum and Shahidi discussed the implications of this problem for estimated measures of income inequality.

Values of long-run supply elasticity implied by the three auxiliary models (yield, herd size, and population) were calculated as the sum of the short-run elasticity of yield, the long-run elasticity of herd size, and the long-run elasticity of population. The implied values of η_i^* were 0.7151, 0.3798, and 0.7820, respectively, for small, medium, and large farms, compared with the estimates of 0.6725, 0.3814, and 0.7585, respectively, obtained from equation (12). The differences are rather small.

In conclusion, it appears that λ_i increases as herd size increases, that negligible herd size adjustments to price changes are accomplished in the short run, and that price increases stimulate little herd size expansion even in the long run on small and medium-sized farms. These findings for small and medium-sized farms are consistent with Kisleev and Peterson's argument that price increases stimulate increased population but not increased herd size and that farm consolidation is caused by factors other than price.

The estimates obtained from the "all farm" data set are useful in evaluating the aggregate

effects of price changes on output via yield, population, and herd size. Because the data set does not involve any grouping, crossover entry and restrictions on herd size do not affect estimates of λ_i and β_i . Results from equation (12) suggest that 71% of all adjustments to a price change are accomplished in the short run ($\eta = 0.3646$), while 29% are accomplished beyond the short run ($\eta^* - \eta = 0.1486 = \lambda_i + \beta_i$). The estimates from the auxiliary models suggest that 63% of the adjustments beyond the short run are herd size adjustments, while 37% are population adjustments.¹⁶ This result contradicts Kisleev and Peterson's finding that price changes do not affect the average farm's size.

The estimates of Y_i and Y_i^* reported in table 4 were generated from the values of η_i and η_i^* obtained via equation (12). The Y_i measures for small, medium, and large-sized farms are, respectively, 0.2532, -0.0007, and -0.0361. The respective measures of Y_i^* are 0.0861, -0.2110, and 0.1661. For illustrative purposes, the effects of a 50% fall in milk prices are shown in table 5. The fall in milk price lowers the output and revenue shares of small farms by 12.66% but raises those of medium and large-sized farms by 0.35% and 1.805%, respectively, in the short run. Consequently, using the 1985 revenue share figures as baseline figures, the output or revenue share of small farmers is lowered from 7.3% to 6.3% in the short run, that of medium-sized

¹⁶ These percentages are obtained by dividing the long-run elasticity of herd size obtained from the yield equation by the sum of the long-run elasticities of herd size and population obtained from the yield and population equations.

farms is virtually unchanged, and that of large farmers is raised from 50.2% to 51.1%. In the short run, price decreases thus seem to help large farms capture a slightly larger share of the market, most of which comes from small farms.

In the long run, a 50% decrease in milk price reduces the revenue or output share of small farms from 7.3% to 6.9%. Thus, while the instantaneous response is a lower revenue share for small farms, the extent of the reduction is lower as time passes. Similarly, a 50% decrease in milk prices lowers the revenue share of large farms from 50.2% to 46.1% in the long run. Thus, while price increases help raise the revenue share of large farms in the short run, after all adjustments are completed, revenue share is actually reduced. The revenue or output share for medium-sized farms is increased from 42.5% to 47.0% in the long run.

The patterns of adjustments observed for the Northeast occur because small and large farms have more elastic long-run supplies than the overall industry supply while medium-sized farms have less elastic supply than the overall industry supply. The lack of flexibility on medium-sized farms helps them capture a larger share of the market in the long run when prices are falling.

The revenue and population data in table 5 are used to construct Lorenz-type curves in figure 1. In constructing figure 1, however, it is im-

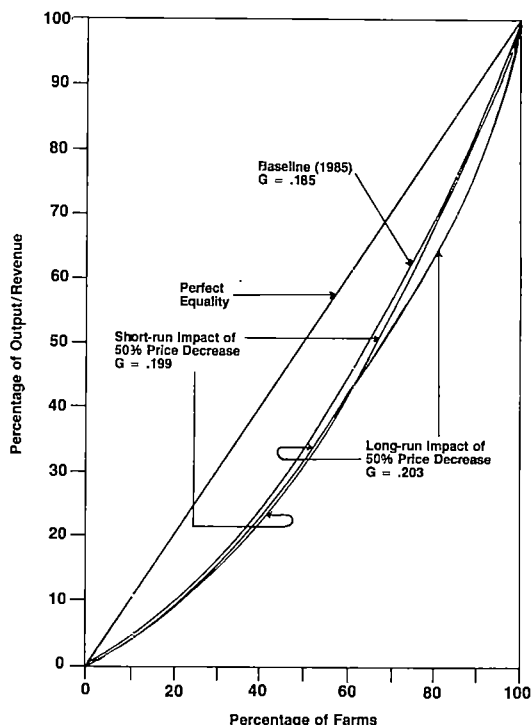


Figure 1. Lorenz curves depicting inequalities in dairy revenue distribution

Table 5. Distribution of Population, Output, and Revenue in Northeastern U.S. Dairy: 1985 and after a 50% Decline in Price

Measure	Herd Size			Gini Coefficient ^a
	Small	Medium	Large	
Population in 1985 and in short run				
Average herd size (milk cows)	33.0	58.0	130.0	
Percentage of farms (b_i)	16.3	48.6	35.1	
Cumulative percentage of farms	16.3	64.9	100.0	
Population in long run				
Percentage of farms (b_i)	16.4	53.8	29.8	
Cumulative percent of farms	16.4	70.2	100.0	
Revenue in 1985				
Percentage of revenue	7.3	42.5	50.2	
Cumulative percentage of revenue	7.3	49.8	100.0	0.185
Revenue in short run				
Percentage of revenue	6.3	42.6	51.1	
Cumulative percentage of revenue	6.3	48.9	100.0	0.199 ^b
Revenue in long run				
Percentage of revenue	6.9	47.0	46.1	
Cumulative percentage of revenue	6.9	53.9	100.0	0.203 ^c

^a Gini coefficient is calculated as $1.0 - \sum_{i=1}^n b_i(z_i + z_{i-1})$, where b_i is the proportion of farms in the herd size interval i and z_i is the proportion of total revenue received by farms in the herd size interval i and all lower intervals (Fichtenbaum and Shahidi).

^b Calculated with b_i evaluated at its 1985 level since $\beta_i = 0$ in the short run.

^c Calculated with b_i values resulting in the long run after population redistribution. The b_i values are based on β_i estimated of .182, .078, and .335, respectively, for small, medium, and large farms.

portant to note that a price change does not only cause output redistribution in the short and long run (through η_i , λ_i , and β_i) but also population redistribution in the long run. Therefore, one must project not only the new output distribution (which Y_i^* helps predict) but also the new population distribution if herd size intervals are held constant.

The elasticity of population share (Θ_i) is obtained through equation (4) as $\Theta_i = \partial \ln b_i / \partial \ln P = \beta_i - \sum_{i=1}^n b_i \beta_i = -0.0032$ for small farms, -0.1072 for medium-sized farms and 0.1498 for large farms. These values are based on the rough estimates of β_i of 0.182 , 0.078 , and 0.335 obtained from the estimation of the population function. Consistent with the assumption that all population redistribution occurs beyond the short run, there is no impact of price on population distribution in the short run (see estimates of population distribution in the short run in table 5). However, the long-run distribution of population will be such that population shares change to 16.4% , 53.8% , and 29.8% , respectively, for small, medium, and large farms, as indicated in table 5. These new population shares were plotted against the revenue shares to obtain the long-run Lorenz curve in figure 1.

The gini coefficient (G) for 1985 is 0.185 . For the situations in the short run and long run after a 50% decline in price, the G s are, respectively, 0.199 and 0.203 (table 5). The diagonal line in figure 1 represents perfect equality of revenue distribution ($G = 0$). Of course, all the Lorenz-type curves are below the perfect equality line ($G > 0$) because farm-level output increases with herd size. Consistent with the gini coefficients, figure 1 shows that a 50% price decrease marginally increases the area between the curve and the perfect equality line in the short run and further increases it in the long run. Reduced price support levels may therefore make revenue distribution more unequal in the short run and even more unequal in the long run. The extent of the revenue redistribution estimated for the northeastern dairy sector is minimal.

Conclusions and Implications

Economists have tended to ignore the effects of price changes on farm population and herd size and the possibility of heterogeneity in the short- and long-run output responsiveness of dairy farms. By conceptualizing and estimating disaggregate measures of supply response, this study

sheds some light on these aspects of price-related policies. The methodology employed in this study is potentially useful for researchers studying other subsectors of U.S. agriculture.

It is theoretically and empirically demonstrated that short- and long-run output responsiveness of dairy farms vary with farm size as a result of size-related differences in the elasticities of yield, herd size, and population. These differences can arise from differences in (a) degrees of capital intensity and specialization, (b) financial ability, and (c) magnitudes of barriers to entry, exit, and crossover. Moreover, differences in output responsiveness imply differentials in the impacts of price changes and price support programs on dairy revenues. The decomposition of the long-run supply elasticity into the elasticities of yield, herd size, and farm population, the conceptualization of the relationships between these elasticities and farm size, and the illustration of resulting revenue redistribution effects of price changes are novel features of this paper.

For the northeastern U.S. dairy sector, it is shown that recent declines in price support levels can create slightly greater inequality in revenue distribution. The limited potential for revenue redistribution in the Northeast does not necessarily suggest that revenue redistribution effects of price changes are always insignificant. Those effects may be significant in subsectors of U.S. agriculture characterized by large differences among farms in supply elasticity. The challenge to researchers is to identify the disparities in supply response before ruling out the possibility of revenue redistribution. In subsectors where significant variations in supply elasticity exist, proposals to change price support levels should consider both the primary beneficiaries and the output, revenue, and perhaps even the profit redistribution effects by employing disaggregate data and the type of analysis conducted in this study.

[Received October 1988; final revision received May 1990.]

References

- Adelaja, A., and K. Rose. "Farm Viability Revisited: A Simultaneous-Equation Cash Flow Approach." *Agr. Finan. Rev.* 48(1988):10-24.
- Antle, J. M. "The Structure of U.S. Agricultural Technology, 1910-78." *Amer. J. Agr. Econ.* 66(1984):414-21.

- Baumol, W. J. "Contestable Markets: An Uprising in the Theory of Industry Structure." *Amer. Econ. Rev.* 72(1982):1-15.
- Baumol, W. J., J. C. Panzar, and R. D. Willig. *Contestable Markets and the Theory of Industry Structure*. San Diego: Harcourt, Brace, and Jovanovich, 1982.
- . "Contestable Markets: An Uprising in the Theory of Industry Structure: Reply." *Amer. Econ. Rev.* 73(1983):491-96.
- Baumol, W. J., and R. D. Willig. "Fixed Costs, Sunk Costs, Entry Barriers, and Sustainability of Monopoly." *Quart. J. Econ.* 96(1981):406-31.
- Binswanger, H. P. "The Measurement of Technical Change Biases with Many Factors of Production." *Amer. Econ. Rev.* 64(1974):964-76.
- Chavas, J.-P., and R. M. Klemme. "Aggregate Milk Supply Response and Investment Behavior on U.S. Dairy Farms." *Amer. J. Agr. Econ.* 68(1986):55-66.
- Chen, D., R. Courtney, and A. Schmitz. "A Polinomial Lag Formulation of Milk Production Response." *Amer. J. Agr. Econ.* 54(1972):77-83.
- Christensen, L. R., and D. W. Jorgenson. "The Measurement of U.S. Real Capital Input, 1929-1969." *Rev. Income and Wealth*, Series 15(1969):292-320.
- Dunne, T., M. J. Roberts, and L. Samuelson. "Patterns of Firm Entry and Exit in U.S. Manufacturing Industries." *Rand J. Econ.* 19(1988):495-515.
- Fichtenbaum, R., and H. Shahidi. "Truncation Bias and the Measurement of Income Inequality." *J. Bus. and Econ. Statist.* 6(1988):335-37.
- Gardner, B. L. "Price Discrimination or Price Stabilization: Debating with Models of U.S. Dairy Policy." *Amer. J. Agr. Econ.* 66(1984):763-68.
- Griliches, A. "A Note on the Serial Correlation Bias in Estimates of Distributed Lags." *Economica* 29(1961):65-73.
- Hall, M., and L. Weiss. "Firm Size and Profitability." *Rev. Econ. and Statist.* 49(1967):319-31.
- Halvorsen, H. W. "Response of Milk Production to Price." *J. Farm Econ.* 40(1958):111-13.
- Hoque, A., and A. Adelaja. "Factor Demand and Returns to Scale in Milk Production: Effects of Price, Substitution and Technology." *Northeast J. Agr. and Resour. Econ.* 13(1984):238-45.
- Hoque, A., A. Adelaja, and P. Ganguly. "Substitution of Energy Inputs, Returns to Scale and Size in Milk Production: A Cost Function Analysis of the Derived Demand for Inputs in a Northeastern Dairy Sample." Pap. presented at the 16th annual meetings of the Atlantic Econ. Soc., Philadelphia, 1983.
- Johnston, J. *Econometric Methods*. New York: McGraw-Hill Book Co., 1972.
- Kessides, I. "Towards a Testable Model of Entry: A Study of the U.S. Manufacturing Industry." Ph.D. thesis, Princeton University, 1982.
- Kinnucan, H. W., and O. D. Forker. "Asymmetry in Farm-Retail Price Transmission for Major Dairy Products." *Amer. J. Agr. Econ.* 69(1987):285-92.
- Kislev, Y., and W. Peterson. "Prices, Technology and Farm Size." *J. Polit. Econ.* 90(1982):578-95.
- Klemme, R., and J.-P. Chavas. "The Effects of Changing Milk Price on Milk Supply and National Dairy Herd Size." *Econ. Issues*, no. 92, Dep. Agr. Econ., University of Wisconsin, June 1985.
- Ladd, G. W., and J. E. Martin. *Application of Distributed Lag and Autocorrelated Error Models to Short-Run Demand Analysis*. Dep. Econ. and Sociology Res. Bull. No. 526, Iowa State University, May 1964.
- Levins, R. A. "Price Specification in Milk Supply Response Analysis." *Amer. J. Agr. Econ.* 64(1982):286-88.
- Matulich, S. C. "Efficiencies in Large-Scale Dairying: Incentives for Future Structural Change." *Amer. J. Agr. Econ.* 60(1978):642-47.
- Milligan, R. A. "Milk Supply Response in California: Effects of Profitability Variables and Regional Characteristics." *West. J. Agr. Econ.* 3(1978):157-64.
- Mills, D. E., and L. Schumann. "Industry Structure with Fluctuating Demand." *Amer. Econ. Rev.* 75(1985):758-67.
- Oi, W. Y. "Slack Capacity: Productive or Wasteful?" *Amer. Econ. Rev.* 71(1981):64-69.
- Panzar, J. C., and R. D. Willig. "Free Entry and the Sustainability of Natural Monopoly." *Bell J. Econ.* 8(1977):1-22.
- Pindyck, R., and D. Rubinfeld. *Econometric Models and Economic Forecasts*. New York: McGraw-Hill Book Co., 1981.
- Salant, P., M. Smale, and W. Saupe. *Farm Viability: Results of the U.S.D.A. Family Farm Surveys*. Washington DC: U.S. Department of Agriculture, Econ. Res. Serv. Rural Develop. Res. Rep. No. 60, 1986.
- Silberberg, E. "The Theory of the Firm in Long-Run Equilibrium." *Amer. Econ. Rev.* 64(1974):734-41.
- Stigler, G. J. "Production and Distribution in the Short-run." *J. Polit. Econ.* 47(1939):305-27.
- Tremblay, R. H. *ELFAC Dairy Farm Business Analysis*, University of Vermont Coop. Extens. Serv. for the Northeastern States Ann. Rep. Series, 1971-1985.
- U.S. Department of Agriculture. *Agricultural Prices*. Washington DC, various issues.
- . *Agricultural Statistics*, Washington DC, various issues.
- U.S. Department of Agriculture, Econ. Res. Serv. *Farm Labor*, Washington DC, various issues.
- U.S. Department of Agriculture, Nat. Agr. Statist. Serv. *Milk Production*. Washington DC, various issues.
- U.S. President, Council of Economic Advisers. *Economic Report of the President, 1987*. Washington DC, 1987.
- Weitzman, M. L. "Constable Markets: An Uprising in the Theory of Industry Structure: Comment." *Amer. Econ. Rev.* 73(1983):486-87.
- Wipf, L., and J. P. Houck. *Milk Supply Response in the United States: An Aggregate Analysis*. Agr. Econ. Rep. No. 532, University of Minnesota, July 1967.
- Wold, H. "On Least Squares Regression with Autocorrelated Variables and Residuals." *Bull. de L'institut International de Statistique* 32(1950):277-89.

Does Arbitraging Matter? Spatial Trade Models and Discriminatory Trade Policies

Giovanni Anania and Alex F. McCalla

When modeling discriminatory trade policies, such as targeted embargoes or targeted subsidies, failure to explicitly include assumptions about the possibility of simultaneous exporting and importing may yield misleading results. Nonlinear programming and "vector sandwich" models implicitly set rules regarding arbitraging which may be at variance with actual policies and/or country behavior. The paper introduces an alternative spatial model which allows the researcher to explicitly incorporate her own assumptions about arbitraging. An analysis of the 1980 U.S. embargo to the USSR shows how the proposed model performs relative to the most frequently used spatial trade models.

Key words: arbitraging, discriminatory trade policies, embargo, spatial models, trade.

The explicit treatment of arbitraging behavior—that is, countries importing and exporting at the same time—in spatial trade models may be necessary for the models to produce valid results. The issue becomes important when analyzing discriminatory national trade policies intended to benefit friends and/or punish enemies. Examples of such policies include the Generalized System of Preferences (GSP), Lomé' Convention preferences, targeted export subsidies, PL 480, selective quotas, and targeted embargoes. These policies create multiple prices and generate possibilities to export and import simultaneously to take advantage of price spreads. Discriminatory trade policies include mechanisms to prevent arbitraging.

In this paper we argue that unless trade models explicitly incorporate the possibility of simultaneously exporting and importing, the choice of the trade model implicitly sets the assumptions on arbitraging. For example, spatial trade

models using reduced-form trade equations exclude the possibility of a country switching from one side of the market to the other as prices change, or of simultaneously exporting and importing.

In the first part of the paper we discuss the role of arbitraging in the design and management of discriminatory agricultural trade policies. The implications for trade policy analysis of different assumptions about arbitraging are briefly addressed. In the second part, the implicit hypotheses about arbitraging associated with two classes of spatial models, nonlinear programming (NLP) models (which include quadratic programming models as a special case) and vector sandwich (VS) models, are discussed.

An alternative model, presented in the third part of the paper, allows countries to switch from one side of the market to the other as prices change and permits the user to incorporate explicit assumptions about arbitraging. The model is an improvement over other spatial trade models when the policies to be analyzed include, for example, a trade liberalization when preferential trade agreements exist, an embargo, or a targeted export subsidy. A numerical example addresses arbitraging behavior associated with the 1980 U.S. embargo against the USSR and shows how the proposed model compares with frequently used models.

Giovanni Anania is an associate professor of agricultural economics, University of Calabria, Italy; Alex F. McCalla is a professor of agricultural economics, University of California, Davis.

Giannini Foundation Paper No. 943.

Giovanni Anania gratefully acknowledges the financial support received from the Italian National Research Council (Consiglio Nazionale delle Ricerche).

The authors wish to thank Mary Bohman, Fabrizio De Filippis, Tu Jarvis, Ila Temu, Quirino Paris, and two anonymous *Journal* referees for their many valuable comments on an earlier draft.

Discriminatory Agricultural Trade Policies and Arbitraging

World agricultural markets contain many discriminatory trade policies in which arbitraging behavior is a matter of concern. Examples include preferential tariffs, targeted export subsidies, embargoes, customs unions, food aid, and preferential import quotas. For example, countries using targeted export subsidies are concerned with preventing targeted importers from reexporting subsidized imports; countries granting preferential market access are concerned with preventing the reexporting of imports from non-preferred countries.

In the United States preferential agricultural tariff reduction agreements include the GSP, the Caribbean Basin Initiative, the Israeli free trade agreement and the recently concluded U.S.–Canada Free Trade Agreement. Other discriminatory trade policies include the Agricultural Trade Development and Assistance Act of 1954 (PL 480), the Export Enhancement Program (EEP), and a series of targeted embargoes including those against the USSR in 1974, 1975, and 1980–81. On preferential imports under its GSP scheme, the United States applies a “rule of origin” which requires that at least 35% (50% of two preferred countries involved) of the value of the article is added in the developing country. The same constraints apply to duty free treatment under the Caribbean Basin Initiative (Organization of American States). The United States–Canada free trade agreement contains a substantial set of “rules of origin” to prevent reexport when different third country tariffs apply in the two countries. PL 480 prevents concessional shipments from disrupting commercial markets (lowering prices) and seeks “commitment from participatory countries that will prevent resale or transshipment to other countries, or use (for other than domestic purposes) of surplus agricultural commodities purchased under the act” (sect. 101). In the early years of PL480, the United States did not apply the principle of additionality (sales must be in addition to commercial demand) to barter sales as it did to sales for local currency (Davis). Strong protests from other exporters about arbitraging of barter sales led to a modification of policy in 1957 (Mortensen, Ezekiel, Kristjanson).

Abbott, Paarlberg, and Sharples state that targeted export subsidies “are often criticized because arbitrage or rerouting of exports in transit can frustrate the subsidizing country’s policies. International trading firms or importing coun-

tries might benefit by reexporting subsidized commodities, as occurred early in the PL 480 program” (p. 724). To enter the EEC under the preferential tariffs granted by the Lomé Convention, exports from the African, Caribbean, and Pacific (ACP) countries must fulfill the conditions stated in Protocol 1 of the convention, concerning the definition of the concept of “originating products.”¹ A similar condition is contained in the EEC’s GSP scheme (EEC). Borrmann, Borrmann, and Stegger (pp. 117–20) argue that the “rules of origin” may have strongly affected the volume of trade generated through the EEC’s GSF scheme. A “country of origin” constraint is also contained in the ASEAN (Association of Southeast Asian Nations) agreement. Koester and Schmitz argue that Kenya imported sugar from the world market and exported it to the EC to capitalize on benefits due under the community preferential import policy. The popular Italian press claimed Israel is exporting frozen orange juice to the EC well in excess of its processing capacity (*La Repubblica*, Dec. 1986).

In all the preferential tariff reduction agreements, the constraints to assure that imports are originating in the beneficiary country are intended to avoid arbitraging and prevent third countries from exploiting the preferential policy. Constraints on the volume of exports receiving preferential treatment are usually included as well.

However, the various “rules of origin” still allow arbitraging. The preferred country can still find it feasible and profitable to import and export at the same time, using low price imports for domestic consumption while exporting domestic production at a higher preferential price. In this case, the quantity arbitrated is implicitly bounded by the “rules of origin” not to exceed domestic consumption.

Finally, a comprehensive study of economic sanctions by Hufbauer and Schott reports that the lack of cooperation of other countries often contributed to sanctions failure. In real-world discriminatory trade policies, arbitraging does matter. Failure to avoid arbitraging may jeopardize the accomplishment of the expected policy goals.

¹ Products originating in the ACP countries are defined, in simple terms, as products wholly obtained in one or more ACP countries, or products which have undergone sufficient working or processing within the ACP countries (third ACP-EEC Convention, signed in Lomé, Togo on 8 Dec. 1984).

Modeling Discriminatory Trade Policies

When discriminatory trade policies are considered, the determination of the net trade positions needs to be based on a spatial trade model, a model capable of reproducing trade flows between each pair of traders. Each region may buy (sell) from (to) different regions at different prices, collecting (paying) different per unit tariffs (subsidies). Any discriminatory trade policy can be equivalently expressed in terms of a tariff or a subsidy. Targeted embargoes are equivalent to country-specific prohibitive export tariffs. Country-specific export (import) quotas may be translated into two export (import) taxes: one, equal to zero, active up to the quota ceiling; the other, prohibitive, active above that ceiling. Food donations represent volume constrained subsidized exports.

Any spatial trade model solution is such that, for each pair of countries, say country i and country j , the domestic prices (p_i and p_j , respectively), must satisfy the following relations (as long as no binding constraint is imposed on the trade flow between the two countries):

$$(1) \quad (p_j - p_i - t_{ij} + \sigma_{ij} - \pi_{ij}) \leq 0;$$

$$(2) \quad (p_j - p_i - t_{ij} + \sigma_{ij} - \pi_{ij})x_{ij} = 0;$$

where t_{ij} is the fixed transportation cost to ship one unit of the commodity from region i to region j , σ_{ij} is the export subsidy that country i pays to its producers for each unit exported to country j , π_{ij} is the import tariff that country j imposes on each unit it imports from country i , and x_{ij} is the nonnegative trade flow from country i to country j . If the trade flow from country i to country j is positive, then the per unit transportation cost plus the tariff minus the subsidy must give the wedge between the two domestic prices. If no shipments occur from country i to country j , then the difference between the two domestic prices must be smaller or, at most, equal to the transportation cost plus the tariff minus the subsidy (implying that shipments from country i to country j are not profitable).

Without interventions, the only possible wedges between domestic prices of trading countries are transportation costs. The matrix of the transportation costs is consistent, which means that the minimum cost path to ship from region i to region j is always the one directly connecting the two regions. In this case, there is no rationale for arbitraging. Finding the market equilibrium solution does not depend on differentiating between transportation costs, subsi-

dies, and tariffs. Given domestic demands and supplies, all that matters is the net sum of the transportation cost plus the import tariff minus the export subsidy for each ordered pair [i.e., pairs (i, j) and (j, i) are different] of countries. This quantity represents a generalized transfer cost. The addition to and subtraction from the transportation costs matrix of nondiscriminatory tariffs and subsidies does not affect its consistency. However, this property may be disrupted by discriminatory trade policies.²

Most agricultural trade models are based on an a priori definition of the sets of the importing and exporting regions. Each country is represented through its excess demand or supply schedule. Thus, the possibility of arbitraging and of a country switching from one side of the market to the other as prices change is assumed away.

When each country's position on the world market is not set a priori, the assumptions about arbitraging are (implicitly) left to the structural characteristics of the specific model.³ These assumptions may strongly affect the solution obtained. For example, in a model with no transportation costs, and each country left free to export and import at the same time, the imposition of a tariff by an importing country on its imports from all but one of the other countries leaves each country's net trade position unchanged. Only trade flows change because the demand of the tariff-imposing country will be satisfied by the exporting countries bypassing the tariff by costlessly rerouting their exports through the preferred country. However, a very different outcome is obtained if regions cannot import and export at the same time.

Discriminatory Trade Policies and Commonly Used Trade Models

Nonlinear Programming Models

The most commonly used spatial trade models are the Quadratic Programming (QP) models developed by Takayama and Judge (Thompson, p. 28) in which an artificial quadratic net quasi-welfare function is maximized subject to a set

² It can vanish as a result of the implementation of preferential tariffs or targeted subsidies. It is lost when country-specific embargoes are imposed. The generalized transfer costs matrix being no longer consistent is a necessary, but not sufficient, condition for arbitraging to be profitable.

³ It should be noted, as pointed out by a reviewer, that spatial trade models, because of their homogenous good assumption, tend to overemphasize the occurrence of arbitraging. Hence, it is wise to use caution when considering commodities for which this assumption appears to be particularly strong.

of linear constraints. The quasi-welfare function is given by the sum of consumers' and producers' surpluses over all the regions considered. Linear demand and supply functions, large countries, and perfect competition in domestic and world markets are assumed (Takayama and Judge 1964 and 1971, Bawden, Takayama). Rowse expanded the QP formulation of the model to include nonlinear demand and supply functions. Thus, the classical QP model is a special case of the more general nonlinear programming (NLP) formulation.

Takayama and Judge (1971, chap. 10) propose a framework to analyze trading when tariffs and subsidies are present and suggest that it can also be used when discriminatory trade policies are active.⁴ They propose two alternative modeling approaches, based on domestic demand and supply functions (Takayama and Judge 1971, chap. 7, 8), and on excess supply/demand functions (Takayama and Judge 1971, chap. 9). They claim that in a large spectrum of cases the two models are equivalent, and that the second one may be more efficient. However, when discriminatory trade policies are considered, the equivalence of the two models may vanish.

The first model, which uses domestic demand and supply functions, leaves each country free to import and export at the same time but implicitly constrains imports to not exceed domestic consumption. The second model leaves the possibility of arbitrage totally free. These results can be easily verified by comparing the Kuhn-Tucker conditions associated with the two models.⁵ When the first model is used and the

constraint on arbitrage is binding, the arbitrage country's consumption and production prices are not equal. The consumption price is linked to the low price in the region(s) where the imports come from, while the production price is linked to the high price in the region(s) exports are shipped to. This implicit constraint on arbitrage reproduces a condition similar to that imposed by the "rules of origin" observed in real-world preferential trade agreements.

When discriminatory trade policies are present, the two models can yield different results. This will be the case when the solution obtained by using the model based on excess demand and supply functions is such that (a) at least one region exports and imports at the same time, and (b) its imports exceed its consumption.

When each country is represented as both a consuming and a producing region, the Kuhn-Tucker conditions of the NLP problem as formulated by Rowse are analogous to those of the QP formulation of the model based on domestic demand and supply functions. Arbitrage is possible, but in each country imports cannot exceed domestic consumption. If each region is a priori defined as an importer or as an exporter, then, a no-arbitrage constraint is implicitly imposed. The assumptions about arbitrage implicit in nonlinear programming models are summarized in table 1.

Whenever the classical QP models or NLP models are used to analyze markets characterized by discriminatory trade policies such that the generalized transfer costs matrix is not consistent, assumptions about arbitrage are implicitly made. Such assumptions may have serious implications for the trade policy analysis.

Vector Sandwich Models

MacKinnon (1975, 1976) proposed a vector sandwich procedure to solve spatial trade equilibrium problems.⁶ The procedure allows for nonlinear demand and supply functions as well as transportation costs. Holland developed a microcomputer program based on MacKinnon's procedure to solve relatively small, single commodity, spatial equilibrium models. Holland's program is capable of handling import and export tariffs, both ad valorem and per unit. Constraints may be imposed on specific flows as well as on individual countries' overall imports or

⁴ "In this example we use the same demand and supply functions and transportation costs as in chapters 7 and 8, but assume that $\pi_{21} = 1$ and $\pi_{31} = 1$. It is not necessary to assume that $\pi_{1j} = \pi_{2j} = \dots = \pi_{nj}$ for all j ; that is, the tariff may be discriminatory" (Takayama and Judge 1971, p. 201).

⁵ For the quantity formulation of the model based on the domestic demand and supply functions the Kuhn-Tucker conditions are given in (7.2.9.d) in Takayama and Judge 1971, p. 133; the analogous conditions for the price formulation of the same model are given in (8.3.7.a) and (8.3.7.b), p. 159 (the equivalence of the price and the quantity formulations is proven in Takayama and Woodland). These conditions imply that, if the domestic price is different from zero, (a) domestic consumption must be equal to the portion of the domestic production which is consumed domestically plus the sum of all the imports from the other countries, and (b) domestic production must be equal to the portion which is consumed domestically plus the sum of all exports to the other regions. Each country may import and export at the same time, with the constraint that in each country imports cannot exceed domestic consumption. When domestic consumption is entirely satisfied through imports, domestic production is entirely exported. For the second model, the one based on the use of excess supply/demand functions, the Kuhn-Tucker conditions are given in (9.1.27.d), p. 182, and in (9.3.4), p. 194. In this case, if the domestic price is different from zero, domestic production minus domestic consumption plus imports minus exports must be equal to zero. Arbitrage is now left totally unconstrained.

⁶ A good introduction to fixed point theory as a tool in finding economic equilibrium solutions is Zangwill and Garcia, chapters 5, 6, and 7.

Table 1. Implicit Assumptions About Arbitraging of the QP, NLP, and VS Models

Model	Each Country is Represented through:	Implicit Assumptions About Arbitraging
Quadratic programming models (Takayama and Judge 1964 and 1971, Bawden, Takayama)	linear domestic demand and supply functions	arbitraging allowed but constrained not to exceed domestic consumption (imports cannot be reexported)
	a continuous linear excess demand and supply function	arbitraging allowed and unconstrained
Nonlinear programming models (Rowse)	nonlinear domestic demand and supply functions	arbitraging allowed but constrained not to exceed domestic consumption (imports cannot be reexported)
	a nonlinear excess demand or supply function	arbitraging not allowed
Vector sandwich models (MacKinnon 1975 and 1976, Holland)	linear or nonlinear domestic demand and supply functions	arbitraging allowed but constrained not to exceed domestic consumption (imports cannot be reexported)
	a linear or nonlinear excess demand or supply function	arbitraging not allowed

exports. However, discriminatory tariffs and subsidies are not explicitly considered, although the user can overcome the problem by providing the generalized transfer costs instead of the transportation costs. Applications of the vector sandwich method in agricultural trade policy analyses include Holland and Sharples, USDA (chap. 11), Haley, and Kahn and Meilke.

If each country is considered as both consuming and producing, arbitraging is allowed but, again, it is constrained not to exceed domestic consumption. Each country may import and export at the same time, but imports may not be reexported (table 1).

The Model

In this section a spatial trade model is presented to analyze settings where discriminatory trade policies make the generalized transfer costs matrix inconsistent. The model has two main features: (a) each country can move from one side of the market to the other as prices change, and (b) the user is allowed to explicitly specify assumptions on arbitraging through two sets of parameters. The model is given in two forms, one based on domestic supply and demand functions, the other based on excess demand/supply functions. These functions need not be linear. The model based on domestic demand and supply functions can be constrained so that arbitraging (a) cannot occur, (b) is allowed but is constrained not to exceed domestic consump-

tion, or (c) is allowed and left completely unconstrained. When using the model in which each country is represented through its excess demand/supply function, only the first and the third scenarios can be implemented (table 2). The main advantage of the model we propose over the commonly used models is that it provides the researcher with the possibility of easily comparing within a single model structure the effects of discriminatory trade policies under different hypotheses about arbitraging.

Only one commodity is considered in the model proposed. In addition, a partial equilibrium framework, fixed exchange rates, and perfect competition on domestic and world markets

Table 2. Assumptions About Arbitraging Which Can Be Explicitly Incorporated in the Model Proposed

Each Country is Represented through:	Assumptions About Arbitraging Which Can Be Explicitly Incorporated
Linear or nonlinear domestic demand and supply functions	Arbitraging allowed and unconstrained
	Arbitraging allowed but constrained not to exceed domestic consumption
	Arbitraging not allowed
A linear or nonlinear continuous excess demand and supply function	Arbitraging allowed and unconstrained
	Arbitraging not allowed

are assumed. The model maximizes an artificial quasi-welfare function (W) defined as in Samuelson, Takayama and Judge, and Rowse. When domestic demand and supply functions are considered, the model may be stated as follows:

$$(3) \quad \max_{x_{ij}} W = \sum_i \theta_i(y_i) - \sum_i \phi_i(s_i) - \sum_{ij} [(t_{ij} + \pi_{ij} - \sigma_{ij})x_{ij}],$$

subject to

$$(4) \quad \xi_i^2 \left\{ [1 - (\xi_i - 1)/-2] \left[\sum_j x_{ji} - x_{ii} \right] + [(\xi_i - 1)/-2] \left[\sum_j x_{ij} - x_{ii} \right] \right\} = 0, \\ i = 1, \dots, n;$$

$$(5) \quad \Psi_i \left(\sum_j x_{ji} - x_{ii} \right) - y_i \leq 0, \quad i = 1, \dots, n;$$

$$(6) \quad x_i = \sum_j x_{ij} - \sum_j x_{ji}, \quad i = 1, \dots, n;$$

$$(7) \quad y_i = \max\{x_{ii}, x_{ii} - x_{ij}\}, \quad i = 1, \dots, n;$$

$$(8) \quad s_i = y_i + x_i, \quad i = 1, \dots, n;$$

$$(9) \quad x_{ij} \geq 0; \quad i, j = 1, \dots, n;$$

where i and j denote the regions ($i, j = 1, 2, \dots, n$); y_i denotes the quantity consumed in country i ; s_i denotes the quantity produced in country i ; $\theta_i(y_i)$ denotes the integral under the inverse domestic demand of region i , $p_i^d(y_i)$, between 0 and y_i ; $\phi_i(s_i)$ denotes the integral under the inverse domestic supply of region i , $p_i^s(s_i)$, between 0 and s_i ; x_{ij} denotes the flow of commodity from region i to region j ; x_i denotes the total exports (if positive) or the total imports with the sign changed (if negative) of region i ; t_{ij} denotes the fixed per unit transportation cost for shipping the commodity from region i to region j ; π_{ij} denotes the per unit tariff imposed by region j on its imports from region i ; σ_{ij} denotes the subsidy paid by region i for each unit exported to region j ; ξ_i denotes a parameter controlling the possibility of the i th region to arbitrage, and, if arbitraging is not allowed, the side of the market on which it appears (this parameter may be set to be equal to -1 , 0 or 1 : it will be equal to 0 for the nonbeneficiary re-

gions, and for the beneficiary ones which are left free to arbitrage; to -1 for the beneficiary countries which are not allowed to arbitrage and may operate on the market as importers only, to 1 for those which are not allowed to arbitrage and may operate on the market as exporters only and Ψ_i denotes a parameter constraining arbitraging, when it is allowed, not to exceed domestic consumption (it will be equal to 1 when country i 's imports must not exceed its domestic consumption, to 0 otherwise).

Constraints (6)–(9) are self-explanatory. When Ψ_i in (5) is set equal to 1 , arbitraging cannot exceed domestic consumption. When Ψ_i is equal to 0 arbitraging is not constrained by (5). Constraint (4) allows the user to impose that region i do not arbitrage. When arbitraging is not allowed, (4) provides the model with the position to be taken in the market (importer/exporter) by each region.⁸

Once the model is solved, equilibrium prices are computed as

$$(10) \quad p_i^d = p_i^d(y_i^*); \quad i = 1, \dots, n;$$

$$(11) \quad p_i^s = p_i^s(s_i^*); \quad i = 1, \dots, n.$$

Each country's producers and consumers welfare is defined (fig. 1) as the area between the inverse demand function and the price (the consumers' price if consumers and producers prices are not equal) line, plus the area between the price (the producers' price if consumers and producers prices are not equal) line and the horizontal axis or the inverse supply function, plus the tariff revenue (which is assumed to be redistributed to consumers and producers as a lump sum transfer):

$$(12) \quad W_i = [\theta_i(y_i^*) - y_i^* p_i^d(y_i^*)] + \left[p_i^s(s_i^*) s_i^* - \int_{\max\{0, s_i^*\}}^{s_i^*} p_i^s(s_i) ds_i \right] + \sum_j \pi_{ji} x_{ji}, \quad i = 1, \dots, n$$

⁷ Nonlinear transportation costs may be easily included. They are assumed to be linear to keep the discussion as close as possible to the standard QP setting.

⁸ In the case of a preferential tariff, for example, each beneficiary country has to impose a prohibitive tariff either on its imports or on its exports. The decision on which kind of tariff to implement implies an explicit policy choice: Imports are taxed when the country wants to make use of the preference. Exports are taxed, to make arbitraging unprofitable, when it finds itself better off by importing. The decision is based on the maximization of the beneficiary country's welfare. In many cases, this choice may be easy, as it is the case when only one country is granted a preferential treatment and it is already exporting prior to the implementation of the preferential tariff. In other cases the choice may not be so obvious (Asia). An importing beneficiary country may reach, for example, a higher welfare by remaining on the importers' side of the market even if the preferential treatment granted would make it possible for it to switch to being an exporter.

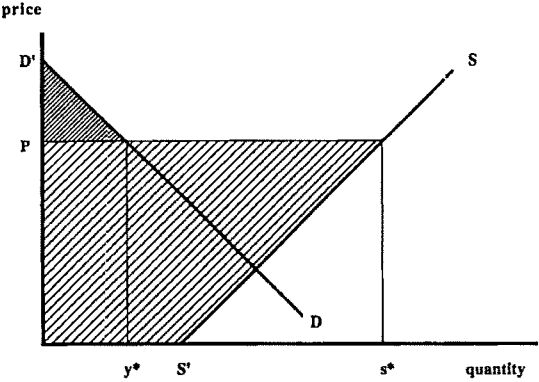


Figure 1. Welfare components for an exporting country

where S'_i is the intercept of the inverse supply function on the horizontal axis. In figure 1 $S'S$ is the inverse supply function, $D'D$ is the inverse demand function, p is the equilibrium price, y^* and s^* are the quantities consumed and produced, respectively, and the crosshatched areas sum to the country's producers' and consumers' welfare.

Often estimates of the domestic supply and demand functions for each region are not available, while estimates of the excess demand/supply functions are. In addition, the excess functions can be more easily estimated. Thus, a model formulation based on excess functions is presented as well. The structure of this model is essentially the same as the one based on the domestic demand and supply schedules. In figure 2 a simple two-country world market case

is presented. S_aS_a and S_bS_b are the inverse excess demand/supply functions of countries A and B, respectively. The market equilibrium is such that region A imports $-X_a$ from region B ($-X_a$ is equal to X_b). The equilibrium prices in the two regions are P_a and P_b , respectively. The difference between the two prices is equal to the transportation cost of shipping one unit from B to A, plus the per unit import tariff imposed by country A, minus the per unit export subsidy paid by country B. The crosshatched areas in figure 2 represent the gains from trade. When excess supply/demand functions are used, W , the artificial quasi-welfare function, may be defined as

(13)

$$W = \sum_i [-\chi_i(x_i)] - \sum_{ij} [(t_{ij} + \pi_{ij} - \sigma_{ij})x_{ij}],$$

where $\chi_i(x_i)$ denotes the integral under the inverse excess supply/demand function of region i , $p_i(x_i)$, between 0 and x_i .

In figure 2 the gains from trade of regions A and B are given by the sum of the areas CDP_a and P_bGF . These gains may be obtained by subtracting the areas FGX_b0 and CP_aP_bE from the area $CDOX_a$. This is given by expression (13) for the n -countries case. $\sum_i [-\chi_i(x_i)]$, in fact, gives the algebraic sum of the areas under each region's excess supply/demand function, as a positive term if the region is importing ($x_i < 0$), negative if it is exporting ($x_i > 0$). The net gains from trade are obtained by subtracting the transportation costs plus the tariff revenues minus the

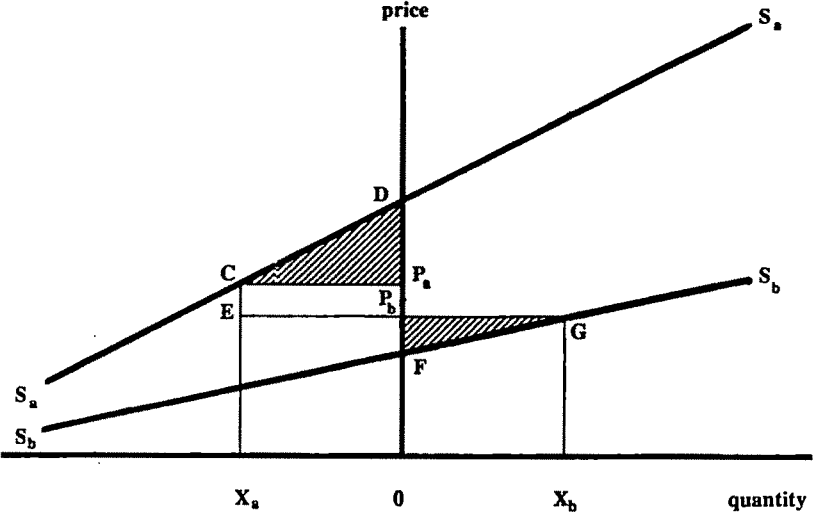


Figure 2. Two-country world trade equilibrium

subsidy expenditure (the algebraic sum of tariffs, subsidies and transportation costs is given in fig. 2 by area CP_aP_bE).

When excess demand/supply functions are used, arbitraging can no longer be constrained to not exceed domestic consumption.⁹

The problem now is stated as

$$(14) \quad \max_{x_{ij}} W,$$

subject to

$$(15) \quad \xi_i^2 \left\{ [1 - (\xi_i - 1)/-2] \sum_j x_{ji} + [(\xi_i - 1)/-2] \sum_j x_{ij} \right\} = 0, \\ i = 1, \dots, n;$$

$$(16) \quad x_i = \sum_j x_{ij} - \sum_j x_{ji}, \quad i = 1, \dots, n;$$

$$(17) \quad x_{ij} \geq 0; \quad i, j = 1, \dots, n.$$

Equilibrium prices and individual countries gains from trade are computed as

$$(18) \quad p_i = p_i(x_i^*), \quad i = 1, \dots, n;$$

$$(19) \quad W_i = [p_i(x_i^*)x_i^*] - \chi_i(x_i^*) + \sum_j (\pi_{ji}x_{ji}^*), \quad i = 1, \dots, n.$$

A Numerical Example

In this section a numerical example illustrates how the results from the proposed model compare with those which may be obtained using other spatial trade models. While the exercise provides some interesting insights about embargoes, its objective does not go beyond showing the effectiveness of the proposed model in analyzing the effects of discriminatory trade policies by explicitly investigating alternative hypotheses regarding arbitraging and cooperation by the other actors involved.

The focus is on the 1980 U.S. embargo of the USSR. It lasted from 4 January 1980 to 24 April 1981 and covered several agricultural products, including wheat, feed grains, soybeans, meat, and dairy products. The embargo was only par-

tial, because the United States fulfilled its commitment to the 1975 U.S.-USSR trade agreement allowing the export to the USSR of 8 million tons of grains in 1979/80 and 1980/81.

The embargo was motivated solely on the basis of a foreign policy concern, as a retaliation to protest the "USSR invasion of Afghanistan." The decision was based on a Central Intelligence Agency (CIA) estimate that the embargo would strongly affect meat consumption in the USSR (USDA). The CIA estimate assumed full cooperation of all other exporters and ignored the possibility of countries arbitraging. The actual short-run impact of the embargo was substantially smaller than expected. It had a very small impact, if any, on meat consumption in the USSR. The USSR strategy was essentially based on (a) replacing imports from the United States by increased imports from other sources, (b) increasing imports of substitute goods, and (c) slightly reducing its stocks.

The main data source is the comprehensive study mandated by the U.S. Congress (USDA). Consistent with the approach followed so far, only one commodity (wheat) is considered in a partial equilibrium framework. Excess supply/demand functions for twenty-seven regions are derived from base net trade positions, prices, and trade elasticities used in USDA. The transportation costs matrix expands on that used in Holland and Sharples.¹⁰ The model is short run in nature; that is, production is fixed and only consumption, stocks, and trade flows change in response to changes in prices. The model time framework is 1980, the only full year the embargo was in place. Domestic as well as trade policies have been incorporated by including price transmission elasticities in the computation of the trade price elasticities (USDA).

The objective is an a priori evaluation of the effects of a zero constraint on wheat exports from the United States to the USSR. Hence, the base solution has no constraints on the trade flows. This solution is then compared with five different scenarios in which the embargo is active and different hypotheses regarding arbitraging and the cooperation of the other countries are assumed. In the first one it is assumed that no country cooperates and arbitraging can occur; that is, the only constraint imposed is the zero constraint on the U.S.-USSR trade flow. The second scenario assumes that all the other ex-

⁹ However, in this formulation of the problem, as well as in the one based on domestic demand and supply functions, each country's arbitraging may be easily constrained not to exceed a specific amount by inserting an ad hoc constraint.

¹⁰ Detailed information regarding the excess supply/demand functions and the transportation costs matrix used are in Anania and McCalla, which can be obtained from the authors.

porters in the base scenario (Canada, the European Community, Oceania, and Argentina) agree not to increase their exports to the USSR above the pre-embargo levels. However, importers are left free to arbitrage. In the third scenario all countries are cooperating; that is, exporters agree not to increase their exports to the USSR, and importers agree not to arbitrage. The fourth and fifth scenarios differ from the second and the third, respectively, only because Argentina is now assumed not to cooperate. (In 1980 Canada, the European Community, and Oceania agreed not to increase their exports to the USSR, even if their actual level of cooperation remains questionable, while Argentina announced that it was not going to cooperate.)

Trade flows and net-trade positions in the base solution are given in table 3. In the pre-embargo scenario the United States exports 33.7 million tons of wheat, and exports to the USSR equal 5.3 million tons. The other net exporters are Canada (17.5 million tons), the European Community (9.1), Oceania (12.2), and Argentina (4.9). Major importers are Japan (5.7 million tons), East Europe (5.4), USSR (14.4), China (12.2), Egypt (5.3), and Middle East (5.3).

Under the first scenario, in which (a) the United States stops its exports to the USSR, and

(b) other countries do not cooperate (table 4), the embargo has negligible effects on the twenty-seven regions' net trade positions. USSR wheat imports from the United States are replaced by increased imports from Canada and, as a result, USSR total wheat imports decline by only 59,000 tons. The United States, in turn, made up for the embargo on its exports to USSR by (a) increasing its exports toward regions they were already exporting to, and (b) exporting 2 million tons to East Europe and half a million tons to Egypt. Essentially, if the United States imposes the embargo with no cooperation from the other market participants, the policy results in a complete failure. Net trade positions remain (almost) unchanged, and only marginal welfare losses arise from increased transportation costs due to the changes in the trade flows.

In the second scenario the U.S. embargo receives full cooperation from all the regions exporting in the base solution (Canada, EC, Argentina, and Oceania). They agree not to increase their exports to the USSR above the base solution levels. However, importers are left free to arbitrage. This scenario actually is very close to the one the United States was trying to reach in 1980. The results of this simulation (table 5) show that obtaining exporters' cooperation does

Table 3. The 1980 U.S. Embargo of the USSR: Base Solution; Trade Flows and Net Trade Positions (million tons)

Destination	Source					Net Trade Positions
	U.S.	Canada	EC	Oceania	Argentina	
Other Western Europe	1.755					-1.755
Japan	5.698					-5.698
South Africa				.011		-.011
East Europe		5.351				-5.351
USSR	5.282	9.090				-14.373
China	9.395			2.792		-12.187
Mexico	.793					-.793
Central America	2.129					-2.129
Brazil	4.786					-4.786
Venezuela	.744					-.744
South America	3.119					-3.119
Sub-Saharan Africa		2.007			.297	-2.305
Nigeria		.495			.510	-1.006
Egypt		.554	4.729			-5.283
North Africa			4.381			-4.381
India				.067		-.067
South Asia					2.748	-2.748
Indonesia				1.505		-1.505
Thailand				.177		-.177
Southeast Asia					1.345	-1.345
East Asia				2.421		-2.421
Middle East				5.275		-5.275
Net trade positions	33.702	17.498	9.110	12.248	4.901	65.211

Table 4. The 1980 U.S. Embargo of the USSR, Scenario No. 1: Embargo Active, No Country Cooperating; Trade Flows and Net Trade Positions (million tons)

Destination	Source					Net Trade Positions
	U. S.	Canada	EC	Oceania	Argentina	
Other Western Europe	1.758					-1.758
Japan	5.698					-5.698
South Africa				.011		-.011
East Europe	2.073	3.255				-5.328
USSR		14.314				-14.314
China	11.941			.264		-12.205
Mexico	.796					-.796
Central America	2.130					-2.130
Brazil	4.787					-4.787
Venezuela	.744					-.744
South America	3.122					-3.122
Sub-Saharan Africa					2.306	-2.306
Nigeria					1.006	-1.006
Egypt	.532		4.746			-5.278
North Africa			4.373			-4.373
India				.068		-.068
South Asia				2.024	.727	-2.752
Indonesia				1.506		-1.506
Thailand				.177		-.177
Southeast Asia				.485	.862	-1.347
East Asia				2.423		-2.423
Middle East				5.284		-5.284
Net trade positions	33.582	17.569	9.119	12.243	4.900	65.167

Table 5. The 1980 U.S. Embargo of the USSR, Scenario No. 2: Embargo Active, All Exporters Cooperating; Trade Flows and Net Trade Positions (million tons)

Destination	Source							Net Trade Positions
	U. S.	Canada	EC	Oceania	Argentina	East Eur.	O. W. Eur.	
Other Western Europe	3.594							-1.762
Japan	5.699							-5.699
South Africa				.011				-.011
East Europe		8.421						-5.347
USSR		9.090				3.074	1.832	-13.996
China	12.000			.228				-12.228
Mexico	.799							-.799
Central America	2.131							-2.131
Brazil	4.789							-4.789
Venezuela	.745							-.745
South America	3.126							-3.126
Sub-Saharan Africa					2.308			-2.308
Nigeria					1.006			-1.006
Egypt	.548		4.734					-5.282
North Africa			4.379					-4.379
India				.069				-.069
South Asia				1.174	1.584			-2.758
Indonesia				1.508				-1.508
Thailand				.177				-.177
Southeast Asia				1.348				-1.348
East Asia				2.426				-2.426
Middle East				5.295				-5.295
Net trade positions	33.430	17.511	9.113	12.236	4.898	-5.347	-1.762	64.940

not guarantee effective results. USSR wheat imports are now predicted to decrease only by 400,000 tons. USSR import price, on the other hand, raises by \$10.48 because of increased transportation costs (table 9). The USSR substitutes for its imports from the United States thanks to arbitrage. It imports 3 and 1.8 million tons from East Europe and from the non-EC western European countries, respectively. Both regions are net importers and arbitrage increased exports from the United States (East Europe) and from Canada (other West Europe). U.S. exports decrease by only 272,000 tons, while price falls by \$1.32. World wheat trade falls only by 271,000 tons. Hence, these simulations suggest that, even if the United States had obtained co-operation from the other exporters, this would not have assured a significant impact of the embargo because of arbitraging. Arbitraging, however, was not considered a relevant issue during the policy design and implementation.

In the third scenario (table 6) all countries cooperate—exporting countries restrict their exports to the USSR not to exceed pre-embargo levels, and importing countries do not arbitrage. The embargo impact is now significant. USSR wheat imports equal only 9.090 million tons, 5.3 million tons below the pre-embargo level. If

Canada, which is the only country exporting to the USSR, does not exploit market power, then USSR import price is \$13.49 lower than the pre-embargo one (table 9). U.S. exports decrease by 2.8 million tons and the export price by \$13.69. World wheat trade decreases by 3.7 million tons. It should be noted that the United States is not the only region paying a price for the U.S. embargo. Canada's exports fall by almost 600,000 tons and its export price by U.S. \$13.49. EC, Argentina, and Oceania all experience lower exports and prices because of the increased competition from U.S. exports in their traditional markets. (The same argument is in Paddock.)

The fourth and fifth scenarios (table 7, 8) show that Argentina not cooperating was a sufficient condition to make the U.S. effort hopeless. In fact, when all exporters but Argentina are constrained to export to the USSR volumes not exceeding the pre-embargo levels, the impact of the embargo is very small. USSR imports decline only by 371,000 tons when arbitraging is allowed (and some arbitrage occurs), and by 382,000 tons when importing countries are assumed to cooperate fully.

The third scenario, in which all exporters cooperate and no arbitraging occurs, likely reflects the one assumed by the CIA when anticipating

Table 6. The 1980 U.S. Embargo of the USSR, Scenario No. 3: Embargo Active, All Countries Cooperating; Trade Flows and Net Trade Positions (million tons)

Destination	Source					Net Trade Positions
	U.S.	Canada	EC	Oceania	Argentina	
Other Western Europe	1.827					-1.827
Japan	5.713					-5.713
South Africa				.013		-.013
East Europe		5.547				-5.547
USSR		9.090				-9.090
China	11.575			1.043		-12.618
Mexico	.850					-.850
Central America	2.146					-2.146
Brazil	4.818					-4.814
Venezuela	.762					-.762
South America	3.184					-3.184
Sub-Saharan Africa		1.412			.928	-2.340
Nigeria		.015			.993	-1.008
Egypt			5.345			-5.345
North Africa		.843	3.651			-4.494
India				.083		-.083
South Asia				1.310	1.553	-2.863
Indonesia				1.531		-1.531
Thailand				.180		-.180
Southeast Asia					1.381	-1.381
East Asia				2.480		-2.480
Middle East				5.480		-5.480
Net trade positions	30.870	16.907	8.996	12.120	4.856	61.501

Table 7. The 1980 U.S. Embargo of the USSR, Scenario No. 4: Embargo Active, All Exporters but Argentina Cooperating; Trade Flows and Net Trade Positions (million tons)

Destination	Source						Net Trade Positions
	U.S.	Canada	EC	Oceania	Argentina	O.W. Eur.	
Other Western Europe	2.008						-1.763
Japan	5.699						-5.699
South Africa				.011			-.011
East Europe		5.347					-5.347
USSR		9.090			4.668	.244	-14.002
China	12.234						-12.234
Mexico	.800						-.800
Central America	2.131						-2.131
Brazil	4.789						-4.789
Venezuela	.746						-.746
South America	3.126						-3.126
Sub-Saharan Africa		2.146			.158		-2.304
Nigeria		.929			.077		-1.005
Egypt			5.282				-5.282
North Africa	.552		3.829				-4.380
India				.067			-.067
South Asia				2.748			-2.748
Indonesia				1.506			-1.506
Thailand				.177			-.177
Southeast Asia				1.345			-1.345
East Asia				2.422			-2.422
Middle East		1.308		3.970			-5.278
Net trade positions	33.392	17.511	9.111	12.247	4.903	-1.763	64.912

Table 8. The 1980 U.S. Embargo of the USSR, Scenario No. 5: Embargo Active, All Countries but Argentina Cooperating; Trade Flows and Net Trade Positions (million tons)

Destination	Source					Net Trade Positions
	U.S.	Canada	EC	Oceania	Argentina	
Other Western Europe	1.764					-1.764
Japan	5.699					-5.699
South Africa				.011		-.011
East Europe	.214	5.128				-5.342
USSR		9.090			4.901	-13.991
China	12.237					-12.237
Mexico	.800					-.800
Central America	2.131					-2.131
Brazil	4.789					-4.789
Venezuela	.746					-.746
South America	3.127					-3.127
Sub-Saharan Africa		2.303				-2.303
Nigeria		1.003			.002	-1.005
Egypt			5.283			-5.283
North Africa	.554		3.827			-4.381
India				.067		-.067
South Asia				2.749		-2.749
Indonesia				1.506		-1.506
Thailand				.177		-.177
Southeast Asia				1.346		-1.346
East Asia	1.312			1.110		-2.422
Middle East				5.279		-5.279
Net trade positions	33.373	17.525	9.110	12.246	4.904	64.908

Table 9. The 1980 U.S. Embargo of the USSR: Changes in Import and Export Prices as a Consequence of the Imposition of the Embargo (US\$ per ton)

	Base Solution	Scenarios				
		1	2	3	4	5
U.S.	163.28	-.58	-1.32	-13.69	-1.50	-1.60
Canada	165.28	+1.62	+.28	-13.49 ^a	+.30	+.60
EC	177.18	+1.02	+.28	-13.49	-.10	.00
Other Western Europe	179.88	-.58	-1.32	-13.69	-1.50	-1.60
Japan	179.88	-.58	-1.32	-13.69	-1.50	-1.60
Oceania	166.98	-.58	-1.32	-13.69	-.20	-.30
South Africa	192.78	-.58	-1.32	-13.69	-.20	-.30
East Europe	181.28	+1.62	+.28	-13.49	+.30	+.60
USSR	183.08	+1.62	+10.48	-13.49 ^a	+10.30	+10.60
China	192.18	-.58	-1.32	-13.69	-1.50	-1.60
Mexico	177.08	-.58	-1.32	-13.69	-1.50	-1.60
Central America	177.08	-.58	-1.32	-13.69	-1.50	-1.60
Brazil	178.28	-.58	-1.32	-13.69	-1.50	-1.60
Argentina	167.98	-.38	-1.12	-13.49	-.30	+.60
Venezuela	179.78	-.58	-1.32	-13.69	-1.50	-1.60
South America	179.78	-.58	-1.32	-13.69	-1.50	-1.60
Sub-Saharan Africa	196.48	-.38	-1.12	-13.49	+.30	+.60
Nigeria	196.48	-.38	-1.12	-13.49	+.30	+.60
Egypt	191.78	+1.02	+.28	-13.49	-.10	.00
North Africa	191.78	+1.02	+.28	-13.49	-.10	.00
India	196.98	-.58	-1.32	-13.69	-.20	-.30
South Asia	194.78	-.38	-1.12	-13.49	.00	-.10
Indonesia	190.18	-.58	-1.32	-13.69	-.20	-.30
Thailand	188.18	-.58	-1.32	-13.69	-.20	-.30
Southeast Asia	194.78	-.58	-1.22	-13.49	.00	-.10
East Asia	190.18	-.58	-1.32	-13.69	-.20	-.30
Middle East	189.98	-.58	-1.32	-13.69	-.20	-.30

Assuming Canada does not exercise market power.

a strong embargo impact. The fourth scenario—on which arbitraging is allowed, Canada, Oceania, and the EC keep their exports at the pre-embargo levels and Argentina does not cooperate—represents the one closest to the actual outcome.

The results of this analysis suggest that embargoes are not effective policy instruments. Rather, embargo effectiveness rests on one of the two following conditions: (a) all countries cooperate (exporters by freezing their exports to the target country, importers by not arbitraging), or (b) all exporters agree to freeze at the pre-embargo levels their exports not only to the target country but to all importing regions. Both conditions are difficult to achieve.

Finally, if all exporters agree to cooperate, they share part of the cost of the embargo because the embargo imposing country's exports will now displace part of their pre-embargo exports. This implies that asking for cooperation in an embargo scheme should be supported by either a reimbursement for the costs or by the guarantee that the embargo-imposing country's exports would not exceed the pre-embargo level minus

the volume exported to the country the embargo is imposed on.¹¹

¹¹ A second numerical example providing additional evidence on how the model proposed performs compared with NLP and VS models is in Anania and McCalla. This example addresses the hypothetical *ex ante* analysis of the granting of a preferential tariff reduction. Four different solutions are discussed. In the first solution—obtained using the QP and the VS models based on domestic demand and supply functions and the model proposed in this paper, all assuming, implicitly or explicitly, that arbitraging is allowed but constrained not to exceed domestic consumption—one of the beneficiary countries switches from being an importer to being an exporter due to the tariff preference it is granted. Arbitraging occurs and the constraint on its volume is binding. The second solution has been obtained using the QP model based on excess demand and supply functions, and the model proposed leaving arbitraging unconstrained. The nonpreferred exporter increases its exports, completely bypassing the discriminatory tariff. Its exports to the preference granting country are rerouted through one of the countries benefiting of the preferential treatment. The welfare of the preferred countries is lower than that occurring under a non-discriminatory tariff. The third and the fourth solutions are obtained using the model proposed imposing no arbitraging and assuming two alternative hypotheses regarding the beneficiary countries' behavior: that they collude, and that each of them makes its choice between being an exporter and being an importer on the basis of its own welfare only. In these two scenarios the preferred countries reach the highest welfare, and the most efficient policy outcomes (in terms of the welfare the preference-granting country has to give up in order to induce a one-unit increase in the welfare of the beneficiary countries) are obtained.

Conclusions

Arbitraging is an important issue in discriminatory trade policy design and implementation. It may turn discriminatory trade policies into ineffective and very costly policy options. Preventing arbitraging is a very difficult task.

In this paper a simple model has been proposed to analyze discriminatory trade policies. It improves upon commonly used spatial trade models by allowing countries to move from one side of the market to the other as equilibrium prices change, and it allows researchers to properly specify their own assumptions about arbitraging and/or to obtain different solutions as a function of different policy constraints or different levels of effectiveness in enforcing such constraints.

The example presented, focusing on the 1980 U.S. embargo of the USSR, provides evidence on the importance of carefully considering arbitraging in the design and implementation of discriminatory trade policies. Different hypotheses regarding arbitraging (or, more often, the apparently "neutral" choice among alternative models thought to be equivalent) may yield contrasting forecasts of the trade policy impact.

[Received August 1989; final revision received March 1990.]

References

- Abbott, P. C., P. L. Paarlberg, and J. A. Sharples. "Targeted Agricultural Export Subsidies and Social Welfare." *Amer. J. Agr. Econ.* 69(1987):723-32.
- Anania, G. "Welfare Implications of a Preferential Tariff Reduction for Agricultural Exports from Less Developed Countries vs. a Generalized Tariff Reduction." *Agr. Econ.* 3(1989):23-48.
- Anania, G., and A. F. McCalla. "Does Arbitraging Matter? Spatial Trade Models and Discriminatory Trade Policies." International Agricultural Trade Research Consortium, Work. Pap. No. 89-3, 1989.
- Bawden, D. L. "A Spatial Equilibrium Model of International Trade." *J. Farm Econ.* 4(1966):862-74.
- Borrmann, A., C. Borrmann, and M. Stegger. *The EC's Generalized System of Preferences*. The Hague: Martinus Nijhoff Publishers, 1981.
- Davis, J. H. "Agricultural Surpluses and Foreign Aid." *Amer. Econ. Rev.* 49(1959):232-41.
- E.E.C. *Practical Guide to the Use of the European Communities' Scheme of Generalized Tariff Preferences*. Luxembourg: Office for Official Publications of the European Communities, 1982.
- Ezekiel, M. "Impact and Implications of Foreign Surplus Disposal on Developed Economies and Foreign Competitors: the International Perspective." *J. Farm Econ.* 42(1960):163-77.
- Haley, S. L. *Evaluation of Export Enhancement, Dollar Depreciation, and Loan Rate Reduction for Wheat*. Washington DC: U.S. Department of Agriculture, Econ. Res. Serv., Agr. and Trade Anal. Div., April 1989.
- Holland, F. D. *GTP: A Microcomputer Program for the Spatial Equilibrium Problem*. Washington DC: U.S. Department of Agriculture, Econ. Res. Serv., Int. Econ. Div., Oct. 1985.
- Holland, F. D., and J. A. Sharples. *World Wheat Trade: Implications for U.S. Exports*. Dep. Agr. Econ. Staff Pap. No. 84-20, Purdue University, 1984.
- Eufbauer, G. C., and J. J. Schott. *Economic Sanctions Reconsidered: History and Current Policy*. Washington DC: Institute for International Economics, 1985.
- Kahn, L., and K. Meilke. "A Preliminary Investigation of the Effects of General and Targeted Export Subsidies on the World Wheat Market." Dep. Agr. Econ. and Bus. Work. Pap. No. 88/13, University of Guelph, Sep. 1988.
- Koester, U., and P. M. Schmitz. "The EC Sugar Market Policy and Developing Countries." *Eur. Rev. Agr. Econ.* 9(1982):183-204.
- Kristjanson, R. L. "Impact of Surplus Disposal on Foreign Competitors and the International Perspective on Surplus Disposal: Discussion." *J. Farm Econ.* 42(1960):108-83.
- MacKinnon, J. C. "An Algorithm for the Generalized Transportation Problem." *Rgnl. Sci. and Urban Econ.* 5(1975):445-54.
- . "A Technique for the Solution of Spatial Equilibrium Models." *J. Fgnl. Sci.* 16(1976):293-307.
- Mortensen, E. "Impact and Implications of Foreign Surplus Disposal on Developed Economies and Foreign Competitors: the Competitors' Perspective." *J. Farm Econ.* 42(1960):1052-62.
- Organization of American States (General Secretariat-CECON). *The United States Generalized System of Preferences. Caribbean Basin Initiative*. Washington DC, 1984.
- Paddock, B. "Estimation of Producers Losses Arising from the Partial Embargo of Grain Exports to the USSR." *Can. J. Agr. Econ.* 31(1983):233-44.
- Rowse, J. "Solving the Generalized Transportation Problem." *Rgnl. Sci. and Urban Econ.* 11(1981):57-68.
- Samuelson, P. A. "Spatial Price Equilibrium and Linear Programming." *Amer. Econ. Rev.* 42(1952):283-303.
- Takayama, T. "International Trade and Mathematical Programming." *Aust. J. Agr. Econ.* 11(1967):36-48.
- Takayama, T., and G. G. Judge. "Equilibrium among Spatially Separated Markets: a Reformulation." *Econometrica* 32(1964):510-24.
- . *Spatial and Temporal Price and Allocation Models*. Amsterdam: North-Holland Publishing Co., 1971.
- Takayama, T., and A. D. Woodland. "Equivalence of Price and Quantity Formulations of Spatial Equilibrium: Purified Duality in Quadratic and Concave Programming." *Econometrica* 38(1970):889-906.
- Thompson, R. L. *A Survey of Recent U.S. Developments*

- in International Agricultural Trade Models*. Washington DC: U.S. Department of Agriculture, Econ. Res. Serv. Bibliographies and Literature of Agriculture No. 21, 1981.
- U.S. Department of Agriculture, Econ. Res. Serv. *Embargoed, Surplus Disposal, and U.S. Agriculture*. Agr. Econ. Rep. No. 564, Washington DC, Dec. 1986.
- Zangwill, W. I., and C. B. Garcia. *Pathways to Solutions, Fixed Points and Equilibria*. Englewood Cliffs NJ: Prentice-Hall, 1981.

Dynamic Adjustment in the Japanese Livestock Industry Under Beef Import Liberalization

Thomas I. Wahl, Dermot J. Hayes, and Gary W. Williams

Dynamic adjustment of Japanese livestock markets is analyzed under two alternative import policies: the 1988 Japanese Beef Market Access Agreement (BMAA) and complete liberalization of Japanese beef import policies. Modeling problems associated with these policies include the change from a quota to a tariff, the incorporation of a meat demand system derived from utility-maximizing behavior into a livestock policy model, complementarity, and changes in the degree of producer supply responsiveness under alternative policies. An econometric model of the Japanese livestock industry is used to simulate and compare the impacts of the two policies from 1988 through 1997.

Key words: trade liberalization.

Japanese beef producers have been protected from the effects of world markets by government price supports and trade policies since the early 1960s. Japanese beef import suppliers, principally the United States and Australia, have made numerous bilateral and multilateral attempts to liberalize the Japanese market. At the same time, the Japanese government has resisted greater access to its beef markets, preferring relatively small increases in the quota both to appease U.S. interests and to minimize opposition from the politically powerful domestic beef producers. Bilateral negotiations between the United States and Japan in 1988, however, resulted in the Beef Market Access Agreement (BMAA), which was intended to liberalize the Japanese beef market substantially over a six-year period.

The primary purpose of this paper is to analyze the likely dynamic impacts of the recent Japanese beef import agreement upon Japanese livestock supply, meat demand, and beef imports. The results are compared with the effects of continuing a slowly increasing beef import quota, as existed under the previous agreement, and complete liberalization of beef imports. The paper also provides some insights into the gen-

eral analysis of markets under a policy-induced structural change. An econometric simulation model of the Japanese livestock and meat industry is utilized to analyze the likely effects of the alternative policies. A number of associated modeling problems are addressed including the incorporation of an almost ideal demand system (AIDS) for meat into a more complete dynamic model of livestock and meat supply and demand, a shift from an import quota to a tariff as required under the BMAA in 1991, and the out-of-sample price responsiveness of livestock producers.

Following a historical overview of the Japanese livestock industry and policies, a model of the Japanese livestock sector is discussed. Next, the results of simulating the liberalization of beef imports are presented. Finally, some research and policy implications are drawn.

The Japanese Livestock Industry and Policy

The evolution of Japan's political economy following World War II has been the primary motivating force behind the generally protectionist Japanese agricultural policies. Concerns about food security have given agricultural producers a tremendous influence upon Japan's food policies (Longworth). Rural heritages, the electoral system, and the Japanese agricultural cooperatives movement have reinforced the political base

The authors are, respectively, an assistant professor, Department of Agricultural Economics, Washington State University; an assistant professor, Department of Economics, Iowa State University; and a professor, Department of Agricultural Economics, Texas A&M University.

Journal Paper No. J-13379 of the Iowa Agriculture and Home Economics Experiment Station, Project No. 2835.

of farmers. Immediately following World War II, the majority of the population lived and worked on small farms. But as Japan's economy recovered and grew, the population began to shift toward urban centers. Despite the shift in population, representation in the Diet (the Japanese Congress) has been not reapportioned. Consequently, the political influence of the rural districts has increased disproportionately, fostering legislation designed to protect the interests of farmers and agriculturally related industries and to maintain food self-sufficiency levels.

In the 1960s, Japanese self-sufficiency ratios for beef, pork, and chicken meat were all above 90%. By 1988, self-sufficiency ratios for these three meats had dropped to 45%, 80%, and 85%, respectively. Japanese domestic beef supplies come from the native Japanese cattle (Wagyu) herd and the domestic dairy cattle herd. Wagyu cattle have traditionally been fed for long periods (up to 30 months) to produce a highly marbled beef that sells for a substantial premium over the less marbled dairy beef.

Thirty-five percent of the beef from Japanese dairy breeds usually grades as medium and 44% usually grades as common. The typical U.S. choice grade grain-fed beef generally falls between these two grades. Grass-fed beef is most nearly equivalent to the common grade. It is unusual for either dairy or imported beef to make the supreme, superior, or excellent grades that are generally reserved for beef from the Wagyu breed (Miyazaki). There is some evidence that Wagyu beef can be viewed as a separate type of meat and not as a perfect substitute for either dairy or imported beef (Hayes, Wahl, and Williams).

The livestock industry in Japan has had one of the highest levels of protection in the world. Historically, an import quota has been the main policy tool to support the domestic cattle industry and encourage beef production. Through the complicated import quota structure, the government has attempted to maintain the established domestic beef target prices. Then, through a fine-tuning mechanism of purchasing and storing or releasing frozen beef from stocks (the beef price stabilization scheme), the government has stabilized the domestic beef price around the target within a politically and socially acceptable range (the upper and lower stabilization prices). The rapidly increasing demand for beef in Japan has forced the government to allow imports to increase over time to keep prices from increasing significantly above the established stabilization range. As a consequence, Japanese domestic beef

prices have been higher and more stable than otherwise might have been the case.

The Japanese dairy industry is protected by a milk price support program designed to ensure a stable, sufficient supply of milk. The pork industry is protected from world competition by a variable import levy and an ad valorem import tariff. The chicken meat industry is relatively unprotected, but it benefits from the substitution effects of the other policies.

The Beef Market Access Agreement

In June 1988, the Japanese signed the BMAA and agreed to begin a phased reduction of their beef import restrictions. The agreement calls for a 60,000 metric ton annual increase in the import quota during the first three years; i.e., the transitional phase of the agreement beginning in 1988. In the post-transitional phase, however, the quota will be removed completely in 1991 and will be replaced by a 70% tariff.¹ The tariff is to be reduced to 60% in 1992 and to 50% in 1993 (table 1).

Prior to the BMAA, the beef import quota consisted of a general quota and a smaller special quota. Only 10% of the general quota was allocated directly to private traders. The Livestock Industry Promotion Corporation (LIPC) controlled the remainder of the general quota, which was allocated to certain Japanese trading companies through both the simultaneous buy/sell (SBS) system and the tender system. The SBS system allowed Japanese buyers to import beef directly from foreign beef exporters. As shown in table 1, the SBS portion of the LIPC quota was relatively small (10%) in comparison to the tender portion (90%). However, during the transitional phase of the BMAA, the majority of the increase in the global quota has occurred in the SBS portion of the LIPC quota.

The tender portion of the LIPC quota was designed to allow the LIPC to dictate the quality and origin of Japanese beef imports. The LIPC tendered licenses for imports of specific grades of specific beef cuts from specific markets. Licensed trading companies would then purchase the beef from exporters in those specific markets. The BMAA called for a phaseout of LIPC involvement in beef imports in 1990.

¹ Additional safeguards for the Japanese allow for an additional 25% tariff in any year of the post-transitional period in which beef imports increase by 20% or more over the previous year.

Table 1. Japanese Beef Import Quota Levels Under the BMAA (metric tons)

Quota	Transitional Phase				Post-Transitional Phase		
	1987	1988	1989	1990	1991	1992	1993
Global	214,000	274,000	334,000	394,000	472,000 ^a	667,360 ^a	680,832 ^a
Special	20,000	25,000	27,000	30,000	0	0	0
General	194,000	249,000	307,000	364,000	0	0	0
Private	19,000	24,000	307,000	36,400	0	0	0
LIPC ^b	174,600	224,100	276,300	327,000	0	0	0
Tender	157,140	156,870	151,965	151,040	0	0	0
SBS ^c	17,460	67,230	124,335	196,560	0	0	0
Tariff (%)	0.25	0.25	0.25	0.25	0.70	0.60	0.50

^a Level of beef imports that invokes an additional 25% tariff as a safeguard measure.

^b LIPC is Livestock Industry Promotion Corporation.

^c SBS is simultaneous buy/sell.

Effects of the BMAA

Although elimination of beef import restrictions could put pressure on the domestic Japanese cattle industry as well as on the domestic hog, poultry, and fish industries, the extent of the pressure depends upon the degree to which beef imports increase, the response of Japanese cattle producers to expected price changes, the degree of substitutability among meats (including fish) in Japanese consumption, and the policy response of the Japanese government.

If the annual increases in the quota during the transitional period of the BMAA are large enough in comparison to the annual growth in Japanese beef demand over that period, imports might not increase to the higher quota levels and prices in Japan would fall toward world market levels. Even if imports reach the quota level during that period, beef imports could decline during the post-transitional period of the BMAA when the quota is dismantled and the import tariff is implemented. On one hand, if the tariff to be implemented is greater than the tariff equivalent of the quota (i.e., the percentage difference between the domestic and the world prices of import-quality beef), then imports will decline when the tariff is imposed. On the other hand, if the tariff is less than the tariff equivalent of the quota, imports will increase. Thus, whether Japanese beef imports increase or decrease in 1991 from the quota level at the end of the first three years of the BMAA depends largely upon the tariff equivalent of the quota in 1990. If the internal price is more than 70% above the world price of beef in 1990, then Japanese beef imports will increase in 1991. Otherwise, the imposition of the 70% tariff in 1991 could lead to a decline in imports.

The Japanese Livestock Industry Model

The fifty-one-equation model of the Japanese livestock industry used for the analysis of Japanese beef import policy contains three simultaneous sectors: the Wagyu cattle, dairy cattle and beef cattle sector; the hog and pork sector and the chicken meat sector.² The cattle and hog sectors contain two main components: (a) live animal supply (breeding herd, slaughter livestock inventories, animals raised, and imports) and slaughter demand and (b) meat supply (production and imports) and meat demand (consumption). The chicken meat sector contains only meat supply (production and imports) and meat demand (consumption). The livestock supply and demand system is estimated by using three-stage least squares (3SLS), and the meat supply and demand system is estimated by using iterated three-stage least squares (IT3SLS).

Japanese Livestock Supply and Demand

The biological restrictions that are inherent in livestock production are incorporated into the live animal supply and demand component. As a result of these biological restrictions, the source of new animals (i.e., the breeding-herd inventories) "drives" the live animal supply components of the model because future production is determined by current decisions about breeding-herd size. For example, the number of animals born in a given year depends upon the size of beginning breeding-herd inventories. Ending inventories of feeder animals depend upon the

² See Wahl for details of the model, estimation procedures, validation: statistics, and the data used.

number of animals born in the current and previous years. And, because the time from birth to slaughter is constant in an annual model, production in the current period is determined by breeding-herd decisions made in previous periods.

In general, these producer decisions can be represented by the following standard theoretical breeding-herd inventory model for cattle:

$$(1) \quad H_t^* = f(\Pi_{t+1}^e, Z_t),$$

where H_t^* is the desired breeding-herd size at the end of period t , Π_{t+1}^e represents the expected profitability of raising cattle in the following period (the expected cattle price deflated by the expected feed price in period $t + 1$), and Z_t represents other variables, such as technical change, that may affect the desired ending herd size in period t .

A standard form of (1) for estimation can be derived by assuming a partial adjustment framework for inventories and adaptive price expectations:

$$(2) \quad H_t - H_{t-1} = \tau(H_t^* - H_{t-1}), \text{ and}$$

$$(3) \quad \Pi_{t+1}^e - \Pi_t^e = \theta(\Pi_t - \Pi_t^e),$$

where τ is the coefficient of adjustment, θ is the coefficient of expectation, and $0 \leq \tau, \theta \leq 1$. Equation (2) suggests that, primarily because of biological restrictions and the cost of adjustment, changes in breeding-herd size take time; i.e., the breeding herd cannot adjust fully in one year to the long-run desired level. Equation (3) assumes that the change in expected profit in period $t + 1$ is proportional to the current error in forecasting.

Assuming a linear form of (1) and then substituting both (2) and (3) into that equation yields the following breeding inventory model for estimation:

$$(4) \quad H_t = \alpha_0 + \alpha_1 \Pi_t + \alpha_2 H_{t-1} - \alpha_2 H_{t-2} + \alpha_4 Z_t - \alpha_5 Z_{t-1},$$

where

$$\alpha_0 = \tau \theta a,$$

$$\alpha_1 = \tau \theta b,$$

$$\alpha_2 = [(1 - \tau) + (1 - \theta)],$$

$$\alpha_3 = (1 - \tau)(1 - \theta),$$

$$\alpha_4 = \tau c,$$

$$\alpha_5 = \tau(1 - \theta)c.$$

Also, a is the intercept, and b and c are the coefficients of the variables in the linear form of (1); all other variables are as previously defined. The value of θ can be calculated directly from the estimated coefficients of equation (4). The remaining parameters (τ , a , b , and c) can be derived by a procedure suggested by Maddala, given a value for θ .

Because the Japanese government has historically maintained price stabilization bands by using import restrictions and buy/sell programs, producers may consider changes in the government support price to be permanent indicators of future price movements. Movements in market prices outside the price stabilization band have been considered to be only temporary indicators of future price movements. Consequently, Π_t in (4) includes both the market price and the announced government support price, which implicitly assumes that producers form a composite forecast by using both the market and government prices (Granger and Newbold, Hayes and Wahl).

Assuming the framework of (4) for breeding-herd inventories, the live animal supply and demand system for a given livestock type i (Wagyu cattle, dairy cattle, or hogs) is modeled by following the specifications of (5) through (8). Unless otherwise noted, all variables are subscripted by t (year).

$$(5) \quad H_i = \alpha_{i1} + \alpha_{i2} P_i^L / K_i + \alpha_{i3} H_{i,t-1} + \alpha_{i4} H_{i,t-2} + \alpha_{i5} Z_i + \alpha_{i6} Z_{i,t-1} + \alpha_{i7} G_i / K_i,$$

$$(6) \quad R_i = q_i H_{i,t-1},$$

$$(7) \quad F_i = \lambda_{i1} + \lambda_{i2} R_i + \lambda_{i3} R_{i,t-1},$$

$$(8) \quad L_i = H_{i,t-1} - H_i + R_i + F_i - F_{i,t-1}.$$

Endogenous variables are H_i , R_i , F_i , L_i ; H_i is the ending breeding inventory of type i (i = Wagyu cattle, dairy cattle, or hogs); P_i^L is the price of livestock type i ; K_i is the cost of feed for i ; Z_i is the exogenous variable affecting breeding-herd size decisions for livestock type i ; G_i is the government support price of i ; R_i is the offspring of i raised; q_i is the annual birth rate of i ; F_i is the ending feeder inventories of i ; L_i is the slaughter of livestock type i ; and α and λ are estimated parameters.

The production cycle for chickens is only five to six weeks. Consequently, chicken inventory and slaughter equations are less meaningful in an annual model than they are in shorter time frame models. Production changes in response to price changes can occur within such a short time interval that counterintuitive price respon-

siveness may be estimated when using annual data. Thus, live chicken supply and demand equations are not included in the model. However, chicken meat supply and consumption components are included and discussed in the following section.

Japanese Meat Supply and Demand

A reduction in the protection of Japanese beef production would have consequences beyond those for the beef sector alone. Changes in the price and quantity of beef demanded in a given period would affect the demand for other meats, which in turn would alter the prices and, therefore, the supply of other meats in subsequent periods. The dynamic supply adjustments occurring in these other livestock markets would produce feedback effects in the beef market. These supply and feedback effects are usually ignored in most demand analyses, including demand system estimation, but are of crucial importance when the likely magnitude of the price changes and the period over which such changes are expected to occur are relatively large.

The reliability of the analytical conclusions also requires internal consistency in the empirical model. For example, the sum of predicted consumer expenditures on individual meat items should equal the predicted expenditures on all meats. At the same time, the income-compensated effect of a change in beef prices on pork demand should equal the income-compensated effect of a change in pork prices on beef demand. Consequently, the meat demand subsystem in the model should conform to the theoretical restrictions of adding up, homogeneity, and Slutsky symmetry.

Incorporating these theoretical restrictions requires a systems approach to demand estimation. The choice of the aggregation level is also of concern. Ideally, a system including the entire food sector, or even all goods, should be specified and estimated. Because data limitations make including the entire food sector impractical if not impossible, the meats group (including fish) is assumed to be separable from other expenditure groups in this study.

The meat expenditure system used in the model follows the linear approximation of the AIDS (LA/AIDS) specification of Deaton and Muellbauer. Consequently, the budgetary share (W) allocated to the i th meat item in the model is estimated as follows:

$$(9) \quad W_i = \gamma_i + \sum_j \gamma_{ij} \ln P_j + \beta_i \ln E/P,$$

where P_j is the price of good j , E is the per capita expenditures on all five meats, and P is a suitable price index. For P , Deaton and Muellbauer suggest the use of Stone's index (defined as $\ln P = \sum W_i \ln P_i$).

The LA/AIDS places no restrictions on cross price elasticities. Although this lack of restrictions may be desirable from a theoretical viewpoint, utilizing a model with elasticities that are contrary to a priori expectations for policy simulation analysis could lead to counterintuitive results. In this sense, the demand system would not be well behaved. In fact, the estimated elasticities (with homogeneity and symmetry imposed) exhibit complementarity between import-quality beef and chicken meat and between pork and chicken meat. These results would lead to a counterintuitive increase in the simulated level of chicken meat consumption in Japan, for example, given a decline in the price of import-quality beef from import liberalization. The approach taken here was to impose net substitutability on the model by using a Bayesian procedure developed by Geweke as implemented by Hayes, Wahl, and Williams.

To implement this procedure, the estimated parameter vector and variance-covariance matrix of the demand system are used to generate a multivariate normal distribution. Repeated draws from this distribution are then made, testing each vector for net substitutability. The posterior distribution is formed from the parameter vectors that satisfy the inequality restriction:

$$(10) \quad \gamma_{ij} \geq -W_j \cdot W_i.$$

The mean of this posterior distribution will then satisfy net substitutability (see Hayes, Wahl, and Williams for more detail).

The price and expenditure elasticities of the Japanese demand system used in the model (presented in table 2) are Hicksian measures with Slutsky symmetry, homogeneity, and net substitutability among meats imposed upon the estimated IT3SLS coefficients. All own-price elasticities are negative, and all compensated cross-price elasticities are positive. The own-price elasticities are relatively large, especially for beef and imply (*ceteris paribus*) that any reduction in the Japanese beef import barrier would lead to a large increase in the quantity demanded of import-quality beef.³ The estimated expenditure

³ The term import-quality beef is used to refer to the aggregate of imported beef and domestic dairy beef.

Table 2. Japanese Meat Demand and Expenditure Elasticities

Price or Expenditure	Wagyu	IQ Beef ^a	Pork	Chicken	Fish
Wagyu	-2.48	0.44	0.71	0.42	0.90
IQ Beef	0.26	-0.98	0.22	0.04	0.46
Pork	0.24	0.12	-0.73	0.05	0.31
Chicken	0.19	0.03	0.07	-0.91	0.62
Fish	0.08	0.07	0.08	0.12	-0.35
Expenditure	0.75	1.51	0.98	1.15	0.92

Note: All elasticities are Hicksian measures calculated from iterated three-stage least squares (IT3SLS) estimates with Slutsky symmetry, homogeneity, and net substitutability imposed. These results differ slightly from those presented in Hayes, Wahl, and Williams in that they use an iterated seemingly unrelated regressions (ITSUR) estimator.

^a IQ beef is import-quality beef, the aggregate of Japanese dairy beef and imported beef.

elasticities indicate that both Wagyu and import-quality beef are luxury goods in Japan. The expenditure elasticity of demand for import-quality beef is also greater than is that for Wagyu beef. This result is somewhat surprising because Wagyu beef is more expensive in Japan than is import-quality beef.

Integrating the Supply and Demand Systems

The full Japanese livestock industry model combines the meat expenditure system with the supply dynamics for Wagyu cattle and beef (*g*), dairy cattle and import-quality beef (*d*), hogs and pork (*p*), chicken meat (*c*), and fish (*f*). As a large-country representation of the industry, the model incorporates Japanese meat prices and import policies. The live animal component of the model consists of equations (5) through (8) for Wagyu cattle, dairy cattle, and hogs. A conceptual representation of the meat component of the model is represented by equations (11) through (25). The full model, estimated parameters, and validation statistics for the full Japanese livestock industry model are available from the authors.

$$(11) \quad S_i = bL_i \quad (\text{for } i = g, d, p)$$

$$(12) \quad S_c = S_c(P_c^m, T)$$

$$(13) \quad D_i = E_i/P_i^m \quad (\text{for } i = g, d, p, c, f)$$

$$(14) \quad E_i = W_i E \quad (\text{for } i = p, c, f)$$

$$(15) \quad E_i = P_i D_i \quad (\text{for } i = g, d)$$

$$(16) \quad W_i = \gamma_i + \sum_j \gamma_{ij} \ln P_j^m + \beta_i \ln(E/P) \quad (\text{for } i, j = g, d, p, c, f)$$

$$(17) \quad \ln P = \sum_i W_i \ln P_i^m \quad (\text{for } i = g, d, p, c, f)$$

$$(18) \quad P_c^m = P_c(P_c^w)$$

$$(19) \quad V_p = P_p^m - P_p^w$$

$$(20) \quad C_i = D_i \cdot N \quad (\text{for } i = g, d, p, c, f)$$

$$(21) \quad C_g = S_g$$

$$(22) \quad C_d = S_d + \bar{M}_d$$

$$(23) \quad M_i^p = C_i - S_i \quad (\text{for } i = p, c)$$

$$(24) \quad M_i^s = M_i^s(P_i^w) \quad (\text{for } i = d, p, c)$$

$$(25) \quad M_i^s = M_i^p \quad (\text{for } i = d, p, c)$$

where *b* is average slaughter weight;

C_i, total consumption of meat *i* (*);

D_i, per capita consumption of meat *i* (*);

E_i, per capita expenditure on meat *i* (*);

E, per capita consumer expenditure on all meats (*);

L_i, slaughter of livestock type *i* (*i* = *g*, *d*, *p*) (*);

\bar{M}_d , import quota on beef (import-quality);

M_i^s, total import supply for meat *i* (*i* = *d*, *p*, *c*) (*);

M_i^p, total import demand for meat *i* (*i* = *p*, *c*) (*);

P_i^m, domestic price of meat *i* (*i* = *g*, *d*, *c*) (*);

P_i^w, world price of meat *i* (*i* = *d*, *p*, *c*) in domestic currency (*);

P_p^m, fixed support price of pork;

P, meat price index (*);

S_i, domestic supply of meat *i* (*);

S_c , domestic supply of chicken meat (*);
 T , time trend or other shift variables affecting annual chicken meat supply;
 V_F , pork variable levy (*);
 W_i , expenditure share of meat i (*);
 N , Japanese population; and
 β, γ , estimated parameters.

The endogenous variables in the system are marked with an asterisk (*) following the definition of each. Livestock-meat price margin equations link the live animal and meat components of the model [$P_i^L = P_i^L(P_i^m)$ for $i = g, d, p$].

Wagyu beef is modeled as a nontraded commodity. Thus, Wagyu beef is affected by beef imports only through the cross-price effects. Imported beef is treated as the equivalent of dairy beef. Pork is subject to a variable import levy, and chicken is free of import restrictions. The domestic prices of meat (P_i^m) are determined by the interaction of domestic meat supply (S_i) and demand (C_i). The prices of beef, pork, and chicken meat supplied to Japan by the world market (P_i^w) are determined by the excess supplies of each meat facing Japan (M_i^s) and the Japanese import demand for each (M_i^D).

Data Sources

The expenditure and price data used to estimate the parameters of the Japanese livestock industry model presented in the previous section are taken from various yearbooks and reports published by the Japanese Ministry of Agriculture, Forestry, and Fisheries (MAFF). These publications include various issues of *Statistical Yearbook*, *Statistics of Meat Marketing*, *Meat Statistics, in Japan*, *Monthly Statistics of Agriculture, Forestry, and Fisheries*, and the *Annual Report on the Family Income and Expenditure Survey*. Expenditures are calculated as the product of price and disappearance (retail basis) for each meat type. Disappearance for Wagyu beef, import-quality beef, pork, and chicken meat are calculated as the sum of production and imports, the data for which are available from the *Statistical Yearbook*. Data for consumer, wholesale, and producer price indexes, family income, retail fish disappearance, and household family size are also available from the *Statistical Yearbook*. Inventory data for live animals, slaughter numbers, slaughter weights, prices paid to farmers, wholesale prices, and wholesale-to-retail conversion factors are also from the *Statistical Yearbook*. Retail prices for pork and

chicken meat are from *Meat Statistics in Japan*. Retail fish price, from the *Annual Report on the Family Income and Expenditure Survey*, is an average of fresh and salted fish price series weighted by the disappearance of each.

Although a retail beef price is available in the *Meat Statistics in Japan*, individual retail prices for Wagyu and dairy beef are not published. Retail Wagyu and dairy beef prices are calculated by multiplying the respective wholesale prices, available in *Statistics of Meat Marketing* and *Meat Statistics in Japan*, by 2.1156. This coefficient is the average ratio of total retail beef expenditures to the sum of wholesale Wagyu beef expenditures and wholesale dairy beef expenditures.

Beef Import Liberalization

To analyze the effects of switching from a quota (i.e., exogenous imports) to a tariff (i.e., endogenous imports) under the BMAA, a series of conditional statements was added to the model to allow beef imports to fluctuate at or below (but not to exceed) the established quota level during the transitional period of the agreement. During the post-transitional period, beef imports are endogenously determined by supply and demand, which was accomplished by normalizing (23) on beef imports and adding an international price linkage equation that included the tariff (τ) to the model:

$$(26) \quad F_d^m = P_d^w(1 + \tau).$$

Beginning in 1991, the model automatically switches from a quota to the tariff specification, with the tariff set at the BMAA levels. If beef imports do not reach the quota level in any year before 1991, however, the model switches to the tariff specification in that year with the 25% tariff currently in place as called for in the agreement.

To allow for changes in the degree of producer price responsiveness following liberalization, the relative weights in the composite producer price forecast used in (5) shift to account for the discontinuation of the government prices. The weight on the government price is reduced to zero, and the weight on the market price is increased to one. For the Wagyu sector, the consequences are an increase in the elasticity of the breeding herd with respect to the market price from 0.008 to 0.10. The elasticity with respect to the market price in the dairy breeding-herd equation increases from 0.047 to 0.14.

Two alternative policies are also considered to provide comparison with the BMAA: (a) complete, immediate liberalization in 1988 and (b) continuation of the 1984 agreement to allow quota increases of only 9,000 metric tons per year. These two alternatives represent the likely upper and lower limits, respectively, of the beef import liberalization effects of the BMAA on the Japanese livestock industry. Complete liberalization is modeled in the same manner as is the BMAA during the post-transitional phase except that the tariff is set to zero for 1988 through 1997. The continuing quota simulation, on the other hand, assumes that the beef import policy prior to the BMAA remains in place. The quota is increased by 9,000 metric tons per year and the beef support prices are set at their 1987 levels. This quota policy is modeled in the same manner as that of the transitional period of the BMAA except that the quota is allowed to increase by only 9,000 metric tons per year through 1997.

Simulation Results

The simulation results suggest that, under the BMAA, Japanese beef imports will reach 394,000 metric tons by 1990, fulfilling the quota, and then increase by 263,000 metric tons in 1991 as a result of the shift to an import tariff (fig.

1).⁴ Beef imports under the BMAA are projected to reach 1.2 million metric tons by 1997. The substantial increase after 1990 in Japanese beef imports represents an opportunity for U.S., Canadian, Australian, and New Zealand beef exporters to significantly increase exports to Japan.

Complete liberalization of beef imports pushes imports to nearly 1.5 million metric tons in 1988 and to over 2 million metric tons by 1997 because of the increased demand for beef by consumers at the substantially lower world beef prices. In the continuing quota simulation, however, beef imports reach only about 300,000 metric tons in 1997.

The Japanese dairy steer carcass price, which also represents the price of imported beef in the domestic market in this model, is directly affected by the alternative import policies. Under the complete liberalization policy, dairy steer carcass price drops dramatically (by almost 68% in the first year) and remains at about that level for the remainder of the forecast period (fig. 2). Under the BMAA, dairy steer carcass price drops somewhat during the transition from a quota to a tariff in 1990–91 and then drops gradually until 1994, after which the tariff is assumed to remain at 50%.

⁴ The predicted increase in beef imports in 1991 appears to be relatively large. In reality, it is likely that infrastructural constraints and the reticence of consumers to alter expenditure patterns may shift some of the additional imports to later years.

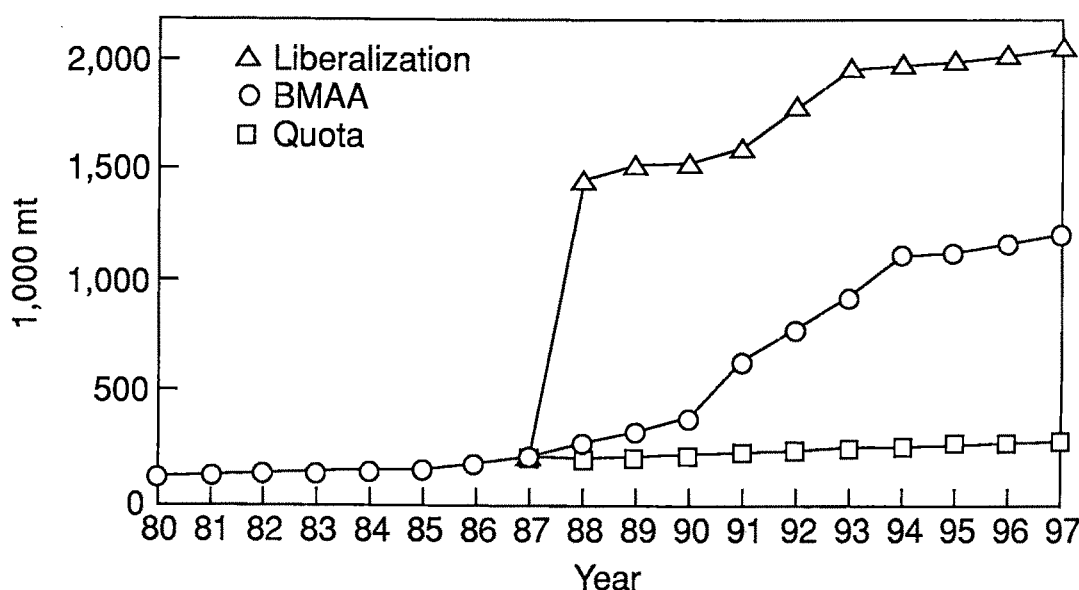


Figure 1. Beef import levels under complete liberalization, the BMAA, and an import quota

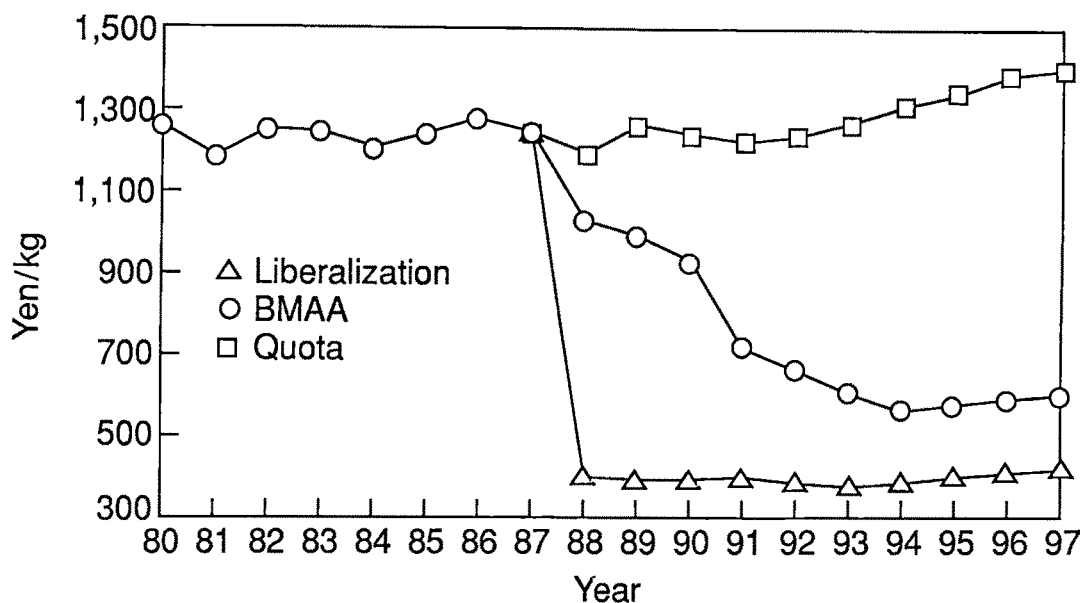


Figure 2. Dairy steer carcass price under complete liberalization, the BMAA, and an import quota

Wagyu steer carcass price (not shown) is affected by the changes in beef import policies only through cross-price effects with the other meats. Under complete liberalization, Wagyu steer carcass price drops by 19% from 1987 to 1988, compared to the 68% drop in dairy steer carcass price. Under the BMAA, Wagyu steer carcass price decreases by less than 4.5% below quota levels until 1990. After 1990, BMAA levels of Wagyu steer carcass price remain about 10% below quota levels.

The Wagyu breeding herd is relatively unaffected by either the BMAA or complete liberalization, falling below projected continuing quota levels by 11% in 1991 and by 5% in 1997 (fig. 3). Initially, the Wagyu breeding herd falls below continuing quota levels under the BMAA and complete liberalization. After 1990, however, the herd rebuilds and follows the same general cycle as the quota. The increase in the breeding herd level after 1988 under the continuing quota results from assumed decreases in feed costs that result from lower grain prices and exchange rates.

Under pressure of growing imports, the dairy cattle breeding herd decreases continuously through the end of the period under the BMAA (fig. 4). Under the continuing quota simulation, however, the dairy cattle breeding herd increases until 1992 because of predicted lower feed costs and continued high milk prices and

then declines gradually until the end of the forecast period. Under complete liberalization, the dairy cattle breeding herd decreases more rapidly than under the BMAA. Nevertheless, by 1997 the complete liberalization level is less than 6% below the BMAA result.

The impact of the alternative import policies on beef production is shown in figures 5 and 6. Wagyu beef supply is nearly equal in the three scenarios by 1997, with little deviation in the period before 1997. The implication is that the BMAA will likely lead to about the same result for the Japanese beef industry as a more immediate, complete liberalization would have, albeit more slowly. Dairy beef supply, on the other hand, is 19% lower under the BMAA and 31% lower under complete liberalization than is the case under the continuing quota scenario by 1997.

Table 3 presents projected dairy beef production, beef imports, and import-quality beef consumption for 1988, 1991, and 1997. Projected per capita consumption for import-quality beef and Wagyu beef and total beef consumption are also presented for comparison. The results suggest that per capita disappearance of import-quality beef under the BMAA could more than double by 1997, whereas Wagyu beef disappearance will likely change little.

The carcass price of hogs changes relatively little over the 1988 to 1997 period under all the policy alternatives because pork import policies

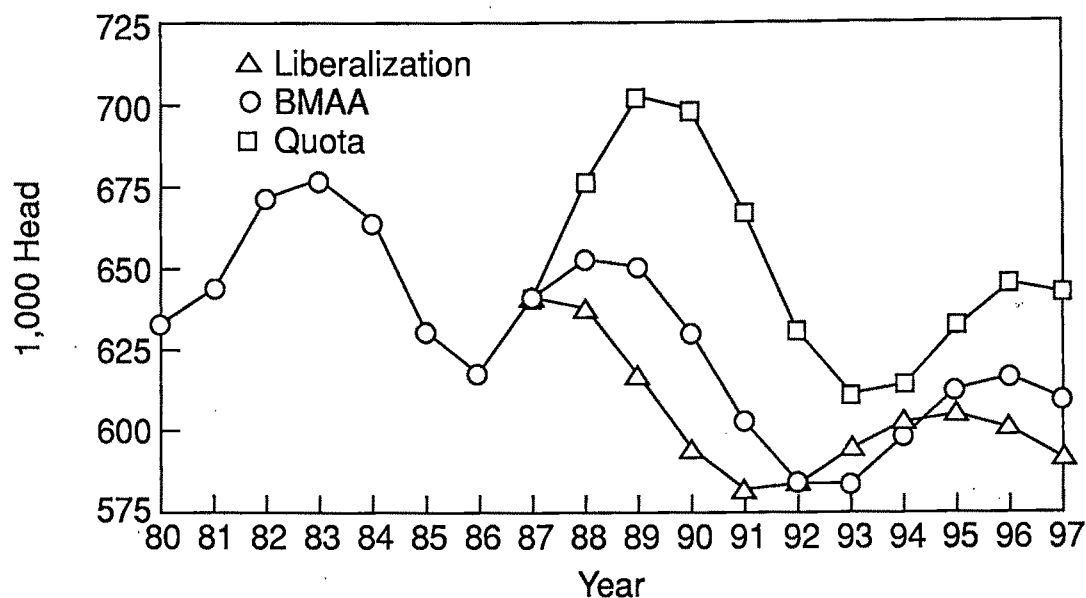


Figure 3. Wagyu cattle breeding-herd ending inventories under complete liberalization, the BMAA, and an import quota

are assumed to remain in place.⁵ Because the changes in hog prices are relatively small, the hog breeding-herd inventory, pork supply, and pork consumption also change little.

Chicken meat supply is projected to increase

⁵ For a discussion of the Japanese pork trade, see Pieri, Meilke, and MacAulay.

at a steady rate under the quota and to remain virtually unchanged under complete liberalization and the BMAA. Chicken meat imports and disappearance under either scenario remain almost unchanged from the levels under the quota. This relative lack of responsiveness to changes in beef import policy results from the low cross-price elasticities in the meat demand system. The

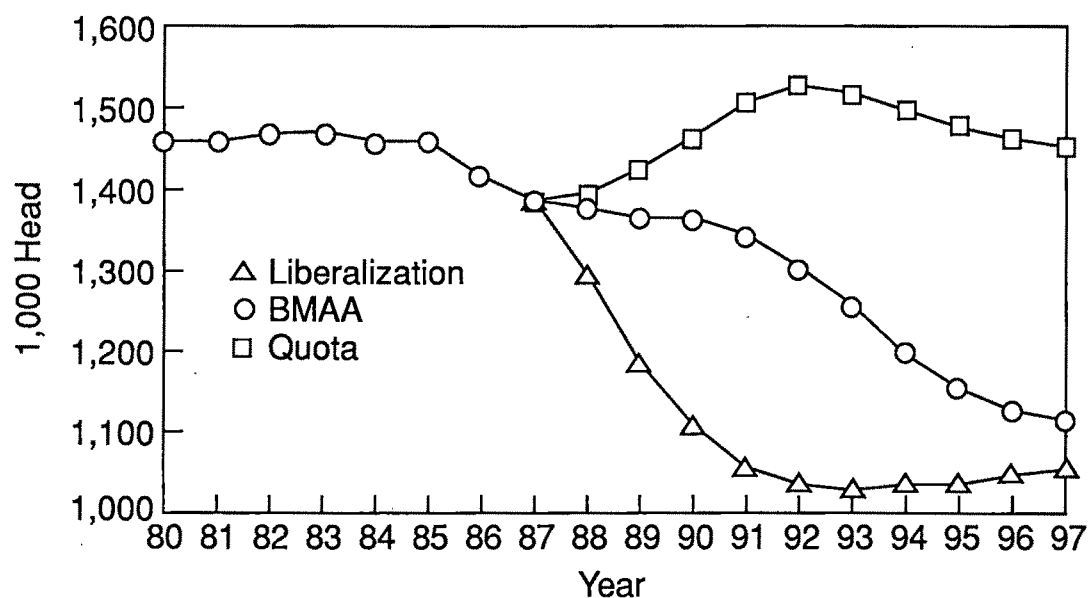


Figure 4. Dairy cattle breeding-herd ending inventories under complete liberalization, the BMAA, and an import quota

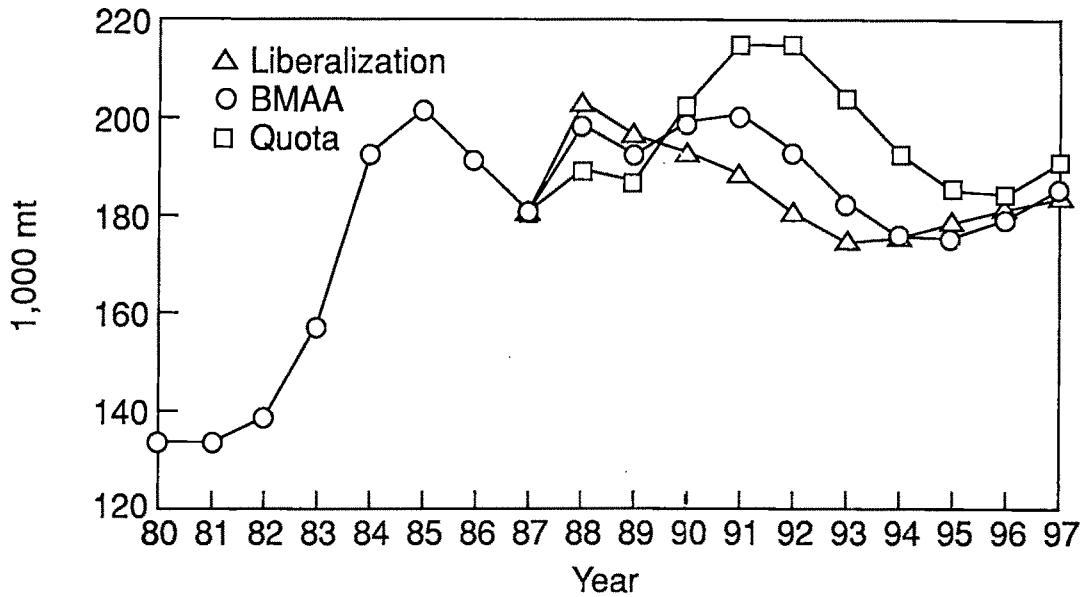


Figure 5. Wagyu beef supply under complete liberalization, the BMAA, and an import quota

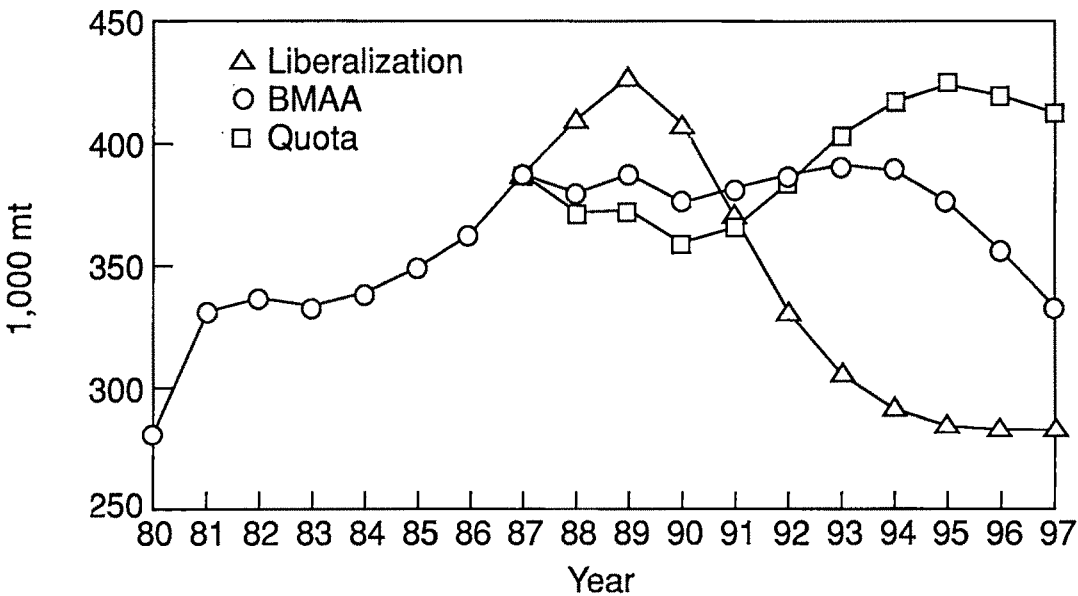


Figure 6. Dairy beef supply under complete liberalization, the BMAA, and an import quota

retail price of chicken meat changes little over the period under the alternative policies.

Sensitivity Analysis

Simulation of the alternative import policy regimes requires a number of assumptions on model

structure and exogenous variables. Two particularly crucial assumptions with significant potential to affect the simulation results involve (a) measuring the tariff equivalent of the current beef import quota and (b) the predicted rate of growth of total consumer expenditures on meat.

Measurement of the price wedge between the domestic and imported beef price is crucial when

Table 3. Projected Japanese Beef Production, Imports, and Consumption for 1988, 1991, and 1997 Under Alternative Beef Import Policies

	Total ^a			Per Capita Consumption ^b		
	Dairy Beef Production	Beef Imports	IQ ^c Beef Consumption	IB Beef	Wagyu Beef	Total Beef
	(1,000 mt)			(kg)		
1988						
Complete liberalization	411	1,470	1,871	10.72	1.15	11.87
MAA	382	274	656	3.86	1.11	4.97
Quota	374	214	588	3.31	1.06	4.37
1991						
Complete liberalization	372	1,610	1,942	10.94	1.05	11.99
MAA	382	657	1,020	5.75	1.11	6.86
Quota	366	241	596	3.36	1.29	4.65
1997						
Complete liberalization	218	2,070	2,300	12.57	0.98	13.55
MAA	332	1,230	1,543	8.43	0.99	9.42
Quota	413	295	694	3.80	1.02	4.82

Carcass basis.

Retail cuts basis.

IQ beef is import-quality beef, the aggregate of Japanese dairy beef and imported beef.

calculating the tariff equivalent of the quota because the magnitude of the price drop imposed under complete liberalization and during the post-transitional phase of the BMAA will have a significant effect on the simulated results. If the rice wedge used to calculate the tariff equivalent is artificially small, understated tariff equivalents will be calculated, and misleading policy implications may result. The appropriate order price to use when calculating the tariff equivalent is the central issue. Unfortunately, accurate historical observations of the border price of beef in Japan are difficult to obtain because the Japanese import several different cuts and

qualities of beef. The Japanese have tended to import more of the high-quality middle cuts to maximize their quota rents. Consequently, the average import unit value is an upwardly biased estimate of the true cost of imported beef.

A comparison of November 1988 U.S. beef prices for chilled and frozen primal cuts and full sets (carcasses) on a CIF Tokyo basis with the corresponding prices of U.S. beef in the Japanese wholesale market is presented in table 4. The calculated tariff equivalent ranges from 63% for frozen tenderloin to 327% for chilled full sets. Note that the tariff equivalents presented in table 4 do not include a profit markup by im-

Table 4. Tariff Equivalents of U.S. Chilled and Frozen Beef, FOB, CIF, and in the Japanese Wholesale Market for November 1988

	U.S. FOB	Insurance and Freight	U.S. CIF	Japan Wholesale	Tariff Equivalent
	(\$/lb)		(Yen/kg)		(%)
Chilled beef					
Striploin	2.88	0.22	918.45	1,900.0	106.40
Tenderloin	4.88	0.22	1,510.99	3,200.0	111.30
Chuck roll	1.41	0.22	482.93	1,950.0	302.87
Full set	1.04	0.22	373.31	1,600.0	327.63
Frozen beef					
Striploin	2.88	0.22	918.45	2,525.0	174.30
Tenderloin	4.88	0.22	1,510.99	2,475.0	63.43
Ribeye roll	3.91	0.22	1,223.61	2,250.0	83.46
Brisket	0.99	0.22	358.49	1,125.0	213.10
Full set	1.04	0.22	373.31	1,190.0	218.05

Note: U.S. prices are taken from the USDA *Market News* and the Japanese wholesale prices of U.S. beef are taken from the *Japan Meat Journal*. The exchange rate used is 134.67 yen/U.S. dollar. The insurance and freight cost is based upon a \$484/mt shipping rate. FOB is the cost free on board and CIF is cost, insurance, and freight.

porters or exporters. The tariff equivalents for chilled beef (except striploin) are much higher than are those for frozen beef, which may reflect either Japanese preference for chilled rather than frozen beef or institutional barriers. Because most of the beef imported by Japan is frozen, the tariff equivalent used in the projections is based on the difference between the frozen full sets price and an assumed 10% profit margin for importers. The resulting tariff equivalent is approximately 190% in 1988 compared to slightly less than 100% when the average import unit value for beef is used to calculate the tariff equivalent.

The potential effects of the alternative measures of the tariff equivalent upon projected BMAA beef import levels are illustrated in figure 7. Using a U.S. price plus transportation costs measure (High T) results in substantially higher levels of beef imports compared to the results of using the average import unit value (Low T). The 400,000 metric tons divergence in the results by 1997 is attributable entirely to the difference in the assumptions about border price levels and demonstrates the importance of accurately measuring the tariff equivalent of the quota.

With regard to total consumer meat expenditures, this study assumes that the annual rate of growth in those expenditures over the forecast period is equal to the 2.5% average annual rate of growth for 1981 through 1987. A higher or lower assumed growth rate would likely have

a significant affect on the forecast growth rate of Japanese beef consumption and, therefore, on beef imports.

Figure 8 illustrates the impact of the BMAA upon beef imports, given a low growth rate (1% and a high growth rate (4%) of meat expenditures, compared to the assumed 2.5% growth rate. The projected level of beef imports in 1999 under these three growth rate assumptions range widely from 1.6 million metric tons under the high growth rate assumption to 1.2 million metric tons under the 2.5% growth rate assumption and 0.9 million metric tons under the low growth rate assumption.

Summary and Conclusions

This paper utilizes dynamic simulation analysis to consider the likely consequences of alternative Japanese beef import policies on the Japanese livestock industry and beef imports. The policies analyzed are the recently signed BMAA complete liberalization, and a continuation of the previous policy of small annual increases in the import quota. Measuring the changes resulting from these policy alternatives created modeling difficulties, which included the incorporation of a meat expenditure system into a more complete dynamic model of livestock supply, a shift from an import quota to a tariff as required under the BMAA in 1991, and the out-of-sample price responsiveness of livestock producers.

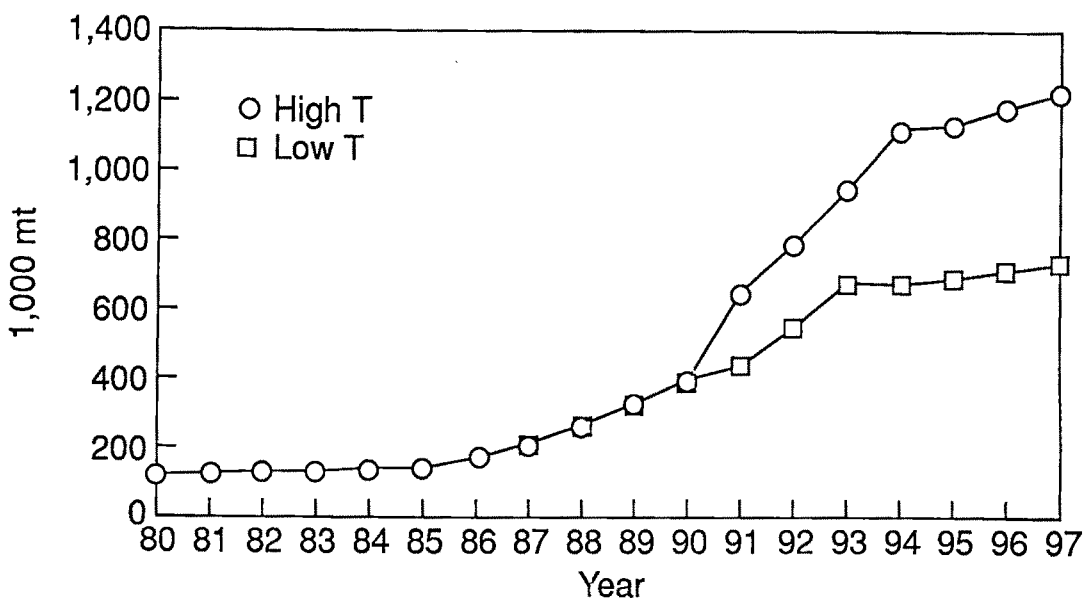


Figure 7. BMAA beef imports under high and low tariff rate assumptions

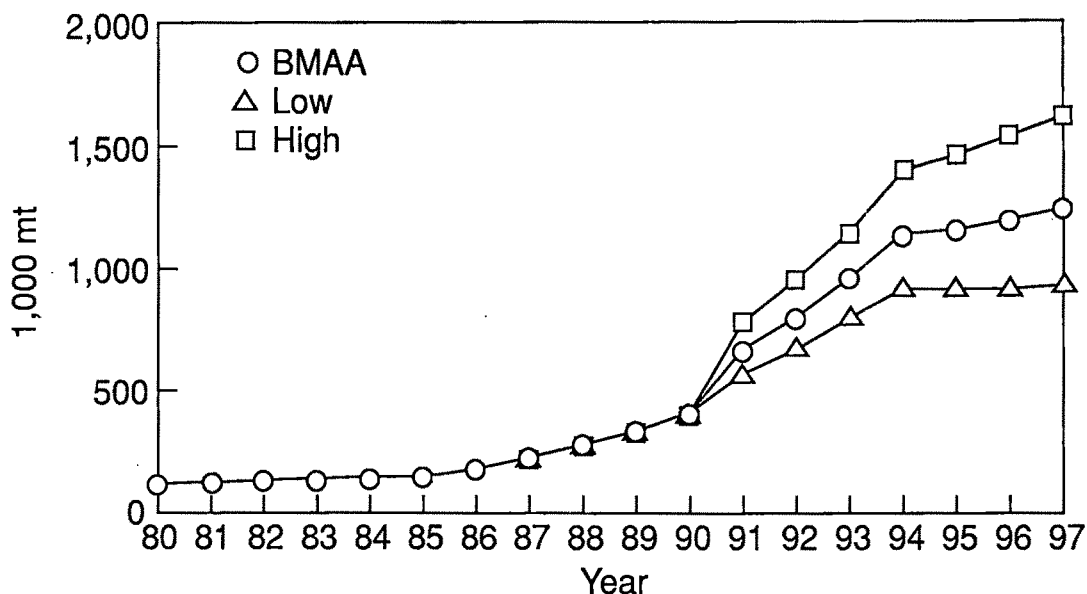


figure 8. BMAA beef imports under alternative total meat expenditure growth rate assumptions

The analytical results lead to a number of conclusions and implications for policy. The projected level of beef imports under complete liberalization suggests that the potential Japanese beef market may be over 1.6 million metric tons by 1991 and over 2 million metric tons by 1997. During the transitional phase of the BMAA, beef imports are likely to increase to the quota level each year. During the post-transitional phase, the additional 25% tariff provision is likely to be invoked. Beef imports are projected to increase by over 67% in 1991 under the BMAA and to reach 1.2 million metric tons by 1997.

The dairy beef industry in Japan will be greatly affected under the policy alternatives. The dairy breeding herd is projected to decrease by about 2% in 1991 under the BMAA and by about 1% under complete liberalization, relative to the results under the quota, because of the substantially lower dairy steer carcass price.

The Wagyu beef industry likely will be affected much less than will the dairy industry under the alternative policies. The Wagyu breeding-herd inventories are projected to decrease within the range of about 11% to 15% in 1991 and from 5% to 8% by 1997, relative to the levels projected under the continuation of the quota, because Wagyu producers are relatively unresponsive to price changes.

The Japanese hog and chicken industries will likely be little affected by the implementation of

alternative policies. The hog industry will be unaffected because the current variable levy is assumed to remain in place. The chicken industry is relatively unaffected because of the low cross-price elasticities with the other meats.

The results of the study are sensitive to assumptions about the tariff equivalent and the growth rates of total meat expenditures used. Consequently, the direction of the change and the comparisons in the movements in the variables under the alternative scenarios in this study, are likely more useful and accurate than are the absolute magnitudes of the simulated values of those variables. Also, many of the modeling techniques developed here may be useful for the other studies of this type that will be necessary during the ongoing GATT negotiations.

[Received December 1988; final revision received March 1990.]

References

- Deaton, A., and J. Muellbauer. "An Almost Ideal Demand System." *Amer. Econ. Rev.* 70(1980):312-26.
- Geweke, J. "Exact Inference in the Inequality Constrained Normal Linear Regression Model." *J. Appl. Econ.* 1(1986):127-41.
- Granger, C. W. J., and P. Newbold. *Forecasting Economic Time Series*, 2nd ed., p. 267. Orlando FL: Academic Press. 1986.

- Hayes, D. J., and T. I. Wahl. "Predicting Changes in the Degree of Producer Responsiveness to Policy Shocks." *NCR-134 Conference: Applied Commodity Price Analysis, Forecasting, and Market Risk Management*. Chicago, 20–21 April 1989.
- Hayes, D. J., T. I. Wahl, and G. W. Williams. "Testing Restrictions on a Model of Japanese Meat Demand." *Amer. J. Agr. Econ.* 72(1990):556–66.
- Japan Meat Journal*, vol. 25, no. 13, Nov. 1988 (in Japanese).
- Japan, Ministry of Agriculture, Forestry, and Fisheries (MAFF). *Annual Report on the Family Income and Expenditure Survey*. Tokyo, various issues.
- . *Meat Statistics in Japan*. Tokyo, various issues.
- . *Monthly Statistics of Agriculture, Forestry, and Fisheries*. Tokyo, various issues.
- . *Statistical Yearbook*. Tokyo, various issues.
- . *Statistics of Meat Marketing*, in Japanese. Tokyo, Japan, various issues.
- Longworth, J. W. *Beef in Japan*. St. Lucia, Australia: University of Queensland Press, 1983.
- Maddala, G. S. *Econometrics*, pp. 144–46. New York: McGraw-Hill Book Co., 1977.
- Miyazaki, A. "Production and Consumption of Beef in Japan." *Research Opportunities in Beef Export Markets United States and Pacific Rim Countries*. Proceeding of an international symposium sponsored by the W 177 Western Regional Beef Research Project and the Farm Foundation, Tucson AZ, 19, 20 Nov. 1986.
- Fieri, K. Meilke, and MacAulay. "North American Japanese Pork Trade: An Application of Quadratic Programming." *Can. J. Agr. Econ.* 25(1977).
- U.S. Department of Agriculture, Agricultural Marketing Service, Livestock Division. *Market News: Weekly Summary and Statistics*, Oct. 1988.
- Wahl, T. I. "Modeling Dynamic Adjustment in Japanese Livestock Markets under Trade Liberalization." Ph.D. thesis, Iowa State University, 1989.

The Competitive Structure of U.S. Agricultural Exports

Daniel H. Pick and Timothy A. Park

The competitive structure of U.S. agricultural exports is examined using a model of exporter behavior based on pricing decisions across destination markets. Market power is revealed in the adjustment patterns of export prices in response to exchange rate movements. The results reject the hypothesis that the export pricing decisions by U.S. firms are consistent with price discrimination across destination markets for cotton, corn, and soybeans. The strongest evidence against the competitive market structure is obtained for international trade in wheat, where results indicate that the two largest importers (Soviet Union and PRC) may exert market power to obtain lower prices.

Key words: agricultural exports, exchange rate, imperfect competition, monopoly, monopsonist, price discrimination.

Recognizing the link between imperfect competition and international trade policy, a diverse set of models of imperfect competition have been developed to explain price formation in international agricultural trade (McCalla, Sarris and Teeb-Tsai, Paarlberg and Abbott, Kolstad and Morris). This paper is motivated by Krugman's observation that tests for imperfect competition in international trade can be based on the observed pricing decisions of exporters. Exporters may exercise market power by adjusting prices to different export destinations, resulting in a form of price discrimination. This pricing-to-market (PTM) behavior pertains to decisions by exporters to maintain or even increase export prices when facing currency depreciation relative to the importer's currency.

The PTM phenomenon has been largely neglected in agricultural trade analysis. Given the dominant U.S. trade shares of many agricultural commodities, pricing decisions by U.S. exporters should be examined for behavior consistent with PTM. For example, the U.S. has been a major exporter of wheat, corn, soybeans, and cotton throughout the 1980s and, until the late 1970s, soybean meal and soybean oil.

In this paper, a model of PTM is applied to U.S. exports of wheat, corn, cotton, soybean, and soybean meal and oil. The main objective

is to develop testable hypotheses about the existence of market power in international commodity trade using the PTM framework. A model of imperfect competition is presented based on pricing decisions by exporting firms across destinations. The model is modified to account for price discrimination by government interventions such as the Export Enhancement Program (EEP).

The first section summarizes previous studies on the competitiveness of international agricultural markets. The second section outlines the basic model and presents an empirical specification designed to test exporter behavior. Sections three and four discuss the data sources, results, estimation, and a model extension. The model extension accounts for the possibility that importers exercise monopsony power in the wheat market. Conclusions and directions for future research then are presented.

Imperfect Competition in Agricultural Markets

The literature review first considers the main approaches to modeling imperfect competition in agricultural commodity trade, mainly in wheat, and then considers models based on industrial organization approaches to imperfect competition in international trade. Imperfectly competitive pricing behavior in the international wheat market was mentioned by Mendelsohn and Farnsworth in the 1950s, but McCalla first in-

Daniel H. Pick is an agricultural economist, Developing Economics Branch, ATAD, Economic Research Service, U.S. Department of Agriculture; Timothy A. Park is an assistant professor, Department of Agricultural Economics, University of Nebraska.

roduced a duopoly model of the international wheat market. His model included Canada and the United States, with Canada as the price leader. Alaouze, Watson, and Sturgess developed a triopoly model which incorporated Australia as a major exporter. Carter and Schmitz proposed a model in which wheat importers exercised monopsony power and imposed import tariffs on wheat to maximize welfare gains. An informal test of the model was based on a comparison of actual prices with empirical estimates derived from the optimal tariff solution.

Oligopoly theory spurred the development of other models investigating imperfect competition in agricultural trade. Sarris and Freebairn modeled international prices as the outcome of a Cournot equilibrium in which pricing policies for individual countries were determined by maximizing domestic welfare. Simulations of the world wheat market were performed to examine the effects of various trade liberalization scenarios. Karp and McCalla applied a Nash cooperative dynamic game to the international corn market and developed multiperiod reaction functions for both importers and exporters. Kolstad and Burris developed a spatial equilibrium model in which countries acted as Nash quantity competitors in the international wheat market. Actual trade flows and predicted trade flows from the model were compared to validate the model.

Paarlberg and Abbott combined a model of oligopoly in international markets with domestic interest group influence to endogenize domestic agricultural policy formation. Model validation was based on comparisons of actual and predicted trade and price levels for wheat.

Two key points should be emphasized. First, price discrimination based on pricing to market and incomplete pass-through of exchange rate movements to export prices has not been addressed for trade in agricultural commodities. Second, the models reviewed here reflect a diverse set of behavioral assumptions, sets of agents and countries, and modeling and validation techniques for imperfect competition in international agricultural trade. The models provide indirect evidence for the validity of their underlying behavioral assumptions using simulation analysis and comparisons of observed and predicted market conditions. Statistical tests for market power which distinguish directly between perfect competition and imperfectly competitive pricing decisions have not been presented.

This paper develops a modeling approach based on firm pricing decisions which yields simple

statistical tests of market power, encompassing perfect competition and two models of imperfect competition. The technique can account for specific market characteristics of a commodity. It is proposed as a prelude to the development of full-scale models of imperfect competition. A prime motivation for this study is to apply the theoretical developments and empirical tests developed in industrial organization to trade in agricultural commodities (Dornbusch, Knetter, Feinberg).

Price Discrimination in International Trade

The scenario developed to identify the competitive structure of a market is based on the incomplete pass-through of changes in the exchange rate to import prices. Krugman discussed the possibility that incomplete pass-through will result from price discrimination by exporters toward importing countries. When export prices of foreign firms are maintained or even increased as the currency of the importing country appreciates, PTM has occurred.

Knetter proposed a model which distinguishes between a competitive market and two models of imperfectly competitive behavior. The exporter is assumed to export to N different markets with individual import demand in each market, $i = 1, \dots, N$, represented as

$$(1) \quad q_{it} = f_i(s_{it}p_{it})v_{it},$$

where q_{it} is the quantity demanded by market i in period t , p_{it} is the export price to market i in the exporter's currency in period t , s_{it} is the i th importer's currency per exporter's currency exchange rate in period t , and v_{it} is a demand shifter. The cost structure for the exporter is a function of the total quantity exported and a cost function shifter δ_t :

$$(2) \quad C_t = C(\sum q_{it})\delta_t.$$

Given (1) and (2), the profit-maximization problem is

$$(3) \quad \text{Max } \pi = \sum_{i=1}^N (p_{it}q_{it}) - C_t.$$

The first-order condition is derived by differentiating (3) with respect to the choice variable p_{it} , and expressed in terms of elasticities:

$$(4) \quad p_{it} = c_t [\epsilon_{it}/(\epsilon_{it} - 1)], \\ i = 1, \dots, N, t = 1, \dots, T;$$

where c_t is the exporter's marginal cost in period t , and ϵ_{it} is the demand elasticity for imports in each of the importing countries in period t .

Expression (4) represents the familiar optimal profit-maximization conditions for the price-discriminating monopolist, equating marginal cost to marginal revenue in each market. When the exporter behaves as a perfect competitor, demand elasticities are infinite and do not vary across destinations. Price then equals marginal cost ($p_{it} = c_t$), and prices are equal across all destinations.

In order to test for alternative market structures, Knetter proposed the following cross-sectional-time-series equation:

$$(5) \quad \ln p_{it} = \theta_t + \lambda_i + \beta_i \ln s_{it} + u_{it},$$

where θ_t is the time effect, λ_i is the country effect, and u_{it} is the error term. Equation (5) can be used to distinguish between three models of market structure. In the competitive market structure, export prices will be the same for all destinations; because there is no country effect, $\lambda = 0$. Changes in the bilateral exchange rates will not affect bilateral export prices, implying $\beta = 0$. The time effects represented by θ_t will measure the common price for all destinations.

The second and third structural models involve imperfect competition with price discrimination across destination markets. The second model assumes constant elasticity of demand with respect to the domestic currency price in each of the importing countries, a reasonable approximation for slight movements along the demand curve. In such a model, the markup over marginal cost as given in (4) is constant but may vary over time and across destinations, implying $\lambda \neq 0$. Shifts in bilateral exchange rates do not influence export prices to various destinations, implying $\beta = 0$.

The third model is based on price discrimination with varying elasticities of demand. Under this scenario, the demand elasticities may vary with changes in the exchange rate. Consider a depreciation of the domestic (importer) currency relative to the foreign (exporter) currency. The price faced by domestic consumers then increases. If the demand elasticities remain constant, then the second case results in which exporters are faced with a constant elasticity demand schedule. However, if demand elasticities change, then the optimal markup over marginal cost will change and export price will thus depend on exchange rates. Krugman referred to this scenario as pricing to market because the

optimal markup by a price-discriminating monopolist will vary across destinations and with changes in bilateral exchange rates. In terms of (5), this model implies that $\lambda \neq 0$ and $\beta \neq 0$.

U.S. Agricultural Exports

The model in equation (5) is applied to five major commodities to test for noncompetitive market structure in United States agricultural exports. The five commodity groups and their respective commodity codes (SITC) included corn (0440045); wheat not donated (0410040); cotton (2631040); soybeans (2222040); and soybean oil, cake, and meal (0810024). Data were compiled on a quarterly basis for the 1978–88 period. Quantity and value data were available from the U.S. Department of Commerce Schedule E. The value data are FAS (free alongside ship), which exclude the cost of loading or any other charges or transportation costs beyond the port of exportation. The quantity and value data were used to generate the price (unit value) variable. Exchange rate data were available from the *International Financial Statistics* published by the International Monetary Fund, while the real exchange rates were calculated using the respective CPI as deflator. Official exchange rates for the Soviet Union, the People's Republic of China (PRC), and Thailand were used in the study.¹

For each commodity a pooled cross-sectional-time-series model was specified with $T \times N$ observations for each model. There are $T - 1$ time dummy effects (θ_t) and $N - 1$ country dummy effects (λ_i). Only the major importers were considered in the analysis.

Tables 1 through 5 summarize the results for each commodity, using both nominal and real exchange rate measures. A significant relationship between export prices in any export destination and the bilateral exchange rate implies a rejection of the constant elasticity model. The number of nonzero coefficients, or violations of the constant elasticity model, was approximately the same for each regression using either nominal exchange rates or real exchange rates.

¹ Although these exchange rates are not determined in the free market, they are adjusted by the respective governments to reflect economic conditions. For the Soviet Union, it is often argued that the price of gold or the trade balance should be used as a proxy to the exchange rate. However, the exchange rate for the Soviet Union is monitored and changed on a monthly basis and is a reasonable proxy for an exchange rate which reflects economic conditions in the Soviet Union. Both exchange rate coefficients for the Soviet Union and the People's Republic of China in the wheat equation were not significant.

Positive and significant coefficients on the exchange rate variable occurred four times for both specifications of the exchange rate variable. A negative coefficient is consistent with a model of price discrimination in which exporting firms adjust prices in export markets to offset local exchange rate movements. Positive coefficients imply that exporters in the United States adjust prices upward as the local currency appreciates, exacerbating the impact of exchange rate movements.

Results for the cotton market in table 1 do not support price discriminating behavior across destination markets. Using the nominal exchange rate measure, only the exchange rate coefficient for Portugal differed significantly from zero at the 10% level. The remaining country effects were not statistically significant. Using the real exchange rate measures produced three significant country effects (Canada, Italy, and South Korea) and no significant exchange rate coefficients. F -tests for $\forall \lambda_i = 0$ and $\forall \beta_i = 0$ were not significant.

Knetter noted that the influence of any particular foreign supplier on prices in a specific

market will be reduced as the number of foreign suppliers to a particular country increases. These results, which indicate the absence of market power, are consistent with the structure of the international cotton market in which a relatively large number of exporters makes it difficult for the United States to exercise market power.

Table 2 summarizes the results of the corn market. The results do not support the hypothesis that the United States as the largest exporter of corn engages in price-discriminating behavior in the international corn market. A possible explanation for the results is that corn is usually exported as feed grain; therefore, the importers face various substitution possibilities among feed grains. Only Mexico, in the nominal exchange rate regression, yielded significant country and exchange rate coefficients. Mexico imports corn for tortilla production and, therefore, corn is imported as food rather than feed grain. The relevant F -tests were not significant.

These results are consistent with Karp and McCalla's model of the international corn market using a noncooperative difference game. The results from the difference game suggested that

Table 1. Country Effects and Exchange Rate Coefficients for Cotton

Destination	Nominal Exchange Rate		Real Exchange Rate	
	λ	β	λ	β
Canada	-0.249 (-1.540)	0.705 (1.392)	-0.293* (-1.670)	1.004 (1.324)
Germany	-0.068 (-0.391)	-0.033 (-0.227)	-0.213 (-1.149)	0.084 (0.555)
Hong Kong	-0.282 (-0.847)	0.023 (0.170)	-0.528 (-1.042)	0.143 (0.524)
Indonesia	-0.500 (-0.936)	0.062 (0.825)	-0.924 (-1.264)	0.120 (1.123)
Italy	0.425 (0.499)	-0.084 (-0.727)	-3.112* (-2.254)	0.418* (2.124)
South Korea	-1.221 (-1.151)	0.163 (1.028)	-2.782* (-1.682)	0.397 (1.582)
Philippines	-0.239 (-1.093)	0.009 (0.124)	-0.502 (0.981)	0.109 (0.495)
Portugal	0.472 (1.450)	-0.117* (-1.799)	-0.482 (-0.597)	0.093 (0.487)
Spain	0.162 (0.316)	-0.075 (-0.705)	0.211 (0.293)	-0.097 (-0.628)
Taiwan	-0.678 (-0.721)	0.130 (0.499)	-1.436 (-1.135)	0.324 (0.945)
Thailand	-0.128 (-0.840)	-0.004 (-0.092)	-0.186 (-1.174)	0.001 (0.029)
United Kingdom		0.076 (0.446)		0.149 (0.755)
	$R^2 = 0.50$		$R^2 = 0.50$	
	$F_{ER} = 0.86$		$F_{ER} = 0.86$	
	$F_{CN} = 1.11$		$F_{CN} = 0.99$	

Note: Values in parentheses are t -values. Asterisks denote t -values significant at the 10% level. F_{ER} is the F -value for testing whether $\forall \beta_i = 0$. F_{CN} is the F -value for testing whether $\forall \lambda_i = 0$.

Table 2. Country Effects and Exchange Rate Coefficients for Corn

Destination	Nominal Exchange Rate		Real Exchange Rate	
	λ	β	λ	β
Belgium	0.012 (0.054)	-0.031 (-0.537)	-0.003 (-0.024)	-0.014 (-0.453)
Japan	-0.229 (-0.687)	0.034 (0.545)	-0.029 (-0.072)	0.005 (0.073)
South Korea	0.242 (0.528)	-0.045 (-0.651)	0.237 (0.324)	-0.037 (-0.331)
Mexico	0.107* (1.680)	-0.021* (-2.464)	0.254 (1.109)	-0.059 (-0.920)
Netherlands	0.046 (0.540)	-0.086 (-1.149)	0.023 (0.266)	-0.006 (-0.089)
Portugal	0.054 (0.382)	-0.030 (-1.063)	-0.279 (-0.772)	0.060 (0.708)
USSR	-0.026 (-0.487)	-0.106 (-1.084)	0.019 (0.318)	-0.094 (-0.952)
Spain	-0.024 (-0.106)	-0.012 (-0.262)	-0.112 (-0.344)	0.018 (0.263)
Taiwan	-0.508 (-1.207)	0.140 (1.208)	-0.293 (-0.515)	0.093 (0.602)
United Kingdom		0.068 (0.909)		-0.324 (-0.214)
	$R^2 = 0.86$ $F_{ER} = 1.18$ $F_{CN} = 0.95$		$R^2 = 0.85$ $F_{ER} = 0.31$ $F_{CN} = 0.31$	

Note: Values in parentheses are *t*-values. Asterisks denote *t*-values significant at the 10% level. F_{ER} is the *F*-value for testing whether $\forall \beta_i = 0$. F_{CN} is the *F*-value for testing whether $\forall \lambda_i = 0$.

U.S. producers gain most when the U.S. behaves competitively even though the United States is in the strongest position to disrupt world trade. By contrast, an oligopoly model developed by Mitchell and Duncan provided weak support for U.S. price leadership in the coarse grain market during the period 1965–81.

The soybean market results are summarized in table 3. The United States is the major exporter of soybeans, accounting for about 70% of global soybean trade. Increased production and exports from Argentina, Brazil, and Paraguay have diminished only slightly the U.S. share. Brazil's trade policy is designed to insure an adequate supply of soybeans for the domestic crushing industry. Trade restrictions and value-added taxes have discouraged the export of soybeans, while tax credits have been used to promote the exports of soybean oil and meal.

The characteristics of international trade in soybeans and processed soybeans suggest that an imperfectly competitive market structure could result. However, the results do not support this hypothesis, mainly because of the use of various substitutes and other exporters for oilseed products. Only in the case of Netherlands, behind Japan as the second largest importer of soybeans from the United States, did the results yield sig-

nificant coefficients for the exchange rate variable (both nominal and real) and the country effect. These significant coefficients for Netherlands may reflect the importance of the crushing industry in the Netherlands. Both *F*-tests were not significant.

The soybean oil, cake, and meal results are summarized in table 4. In the soybean meal market, the U.S. share of global exports is slightly more than 20%. Increased competition from Argentina, Brazil, and the European Community (EC), primarily Spain and Portugal, have steadily eroded the U.S. share since the late 1970s. The U.S. share of global soybean oil exports is similar to that for soybean meal. Argentina, Brazil, and the EC are again the main competitors.

The country effect and exchange rate coefficient for Canada and West Germany were significant when using both the nominal and real exchange rate measures. The joint *F*-tests for the country effects and exchange rate coefficients were significantly different from zero when the nominal exchange rates were used in the estimation. In the specification using real exchange rates, only the *F*-test for the exchange rate coefficients was significant.

The evidence in support of price discrimina-

Table 3. Country Effects and Exchange Rate Coefficients for Soybeans

Destination	Nominal Exchange Rate		Real Exchange Rate	
	λ	β	λ	β
Belgium	-0.199 (-0.962)	0.052 (0.934)	-0.066 (-0.715)	0.017 (0.659)
Germany	0.019 (1.137)	0.008 (0.100)	0.020 (1.166)	-0.002 (-0.031)
Italy	-0.008 (-0.126)	-0.027 (-0.556)	0.001 (0.009)	-0.041 (-0.529)
Japan	0.199 (0.568)	-0.013 (-0.253)	0.289 (0.538)	-0.015 (-0.240)
United Kingdom	0.068 (0.244)	-0.090 (-1.407)	0.081 (0.234)	-0.096 (-1.303)
South Korea	-0.033 (-0.877)	0.002 (0.667)	-0.041 (-0.890)	-0.098 (-0.974)
Mexico	0.022 (1.133)	-0.001 (-0.315)	0.652 (0.992)	-0.026 (-0.466)
Netherlands	0.159* (2.602)	-0.160* (-2.439)	0.154* (2.417)	-0.145* (-2.247)
Spain	0.148 (0.744)	-0.029 (-0.685)	0.107 (0.563)	0.003 (0.690)
Taiwan		0.002 (0.453)		0.002 (0.485)
	$R^2 = 0.85$ $F_{ER} = 0.99$ $F_{CN} = 1.26$		$R^2 = 0.85$ $F_{ER} = 0.78$ $F_{CN} = 1.01$	

Note: Values in parentheses are t -values. Asterisks denote t -values significant at the 10% level. F_{ER} is the F -value for testing whether $\forall \beta_i = 0$. F_{CN} is the F -value for testing whether $\forall \lambda_i = 0$.

ion for soybean oil, cake, and meal was mixed. The significant coefficients on country effects for Canada and West Germany indicate that the United States may have exercised market power in these countries. The findings should be interpreted with caution, however, because the coefficients for the other four countries were not

significant. The results for the soybean meal market reflect the availability of substitutes that exist in the meal market. The availability of substitutes can restrain countries (or firms) from exercising market power.

The strongest evidence to support price discrimination occurred for the wheat market, pre-

Table 4. Country Effects and Exchange Rate Coefficients for Soybean Meal

Destination	Nominal Exchange Rate		Real Exchange Rate	
	λ	β	λ	β
Canada	-0.306* (-2.789)	0.599* (1.687)	-0.342* (-2.072)	1.117* (2.094)
Germany	-0.364* (-2.699)	0.234* (1.819)	-0.310* (-1.798)	0.164 (1.473)
Italy	0.572 (0.987)	-0.109 (-1.330)	1.202 (1.218)	-0.203 (-1.449)
Mexico	-0.133 (-1.329)	0.016 (1.138)	-0.425 (-1.196)	0.109 (1.090)
Netherlands	-0.223 (-1.620)	0.033 (0.282)	-0.184 (-1.043)	-0.007 (-0.066)
Venezuela		-0.047 (-0.979)		-0.050 (-0.560)
	$R^2 = 0.57$ $F_{ER} = 2.13*$ $F_{CN} = 2.68*$		$R^2 = 0.57$ $F_{ER} = 1.93*$ $F_{CN} = 1.78$	

Note: Values in parentheses are t -values. Asterisks denote t -values significant at the 10% level. F_{ER} is the F -value for testing whether $\forall \beta_i = 0$. F_{CN} is the F -value for testing whether $\forall \lambda_i = 0$.

Table 5. Country Effects and Exchange Rate Coefficients for Wheat

Destination	Nominal Exchange Rate		Real Exchange Rate	
	λ	β	λ	β
China	0.187* (2.562)	-0.017 (-0.494)	0.874* (4.525)	-0.011 (-0.334)
Egypt	0.198* (2.459)	-0.024 (-0.335)	0.791* (4.189)	-0.018* (-3.879)
Japan	0.088 (0.274)	0.033 (0.569)	0.885* (2.345)	0.013 (0.206)
South Korea	0.918* (2.397)	-0.107* (-1.827)	1.928* (3.164)	-0.157* (-1.710)
Venezuela	0.237* (2.831)	0.067* (2.404)	0.777* (3.792)	0.165* (3.327)
USSR	0.247* (3.358)	-0.036 (-0.395)	0.934* (4.823)	-0.024 (-0.266)
Taiwan	0.618 (1.604)	-0.089 (-0.844)	1.305* (2.708)	-0.088 (-0.697)
Philippines		0.138* (4.910)		0.472* (5.345)
EEP	-0.213* (-9.317)		-0.237* (-10.981)	
	$R^2 = 0.90$ $F_{ER} = 5.02^*$ $F_{CN} = 2.71^*$		$R^2 = 0.90$ $F_{ER} = 6.746^*$ $F_{CN} = 5.406^*$	

Note: Values in parentheses are t -values. Asterisks denote t -values significant at the 10% level. F_{ER} is the F -value for testing whether $\forall \beta_i = 0$. F_{CN} is the F -value for testing whether $\forall \lambda_i = 0$. EEP represents the coefficient measuring the effect of the EEP.

sented in table 5. Since the analysis includes the period in which the EEP was implemented, equation (5) was modified to account for this policy intervention. The EEP program was established in response to export subsidies by the European Economic Community. Under the program, targeted countries are eligible for subsidized wheat exports. Among the targeted countries included in this study are the People's Republic of China, the Soviet Union, Egypt, and the Philippines. A dummy variable was set equal to one for all periods and countries during which the EEP was in effect and zero otherwise.

The results indicate that five of the country effects and three of the exchange rate coefficients were significant in the regression using the nominal exchange rate. All country effects and four of the exchange rate coefficients were significant when real exchange rates were used. The F -tests of $\forall \beta_i = 0$ and $\forall \lambda_i = 0$ were strongly rejected, thus rejecting the hypothesis of perfectly competitive market structure in the wheat market.

In both regressions the coefficients isolating the EEP effect were negative and highly significant. This result indicates that the EEP had a significant negative impact on the unit export value of wheat to the targeted countries. The impact of targeted export subsidies such as the EEP for the world wheat market was examined

in more details by Seitzinger and Paarlberg using a Nash bargaining model linked to a spatial price equilibrium model.

Wheat export behavior by the United States is not consistent with complete pass-through of exchange rates changes to importers. The evidence supports the model of imperfect competition with price discrimination across destination markets. Pick and Skully confirmed that significant premiums and discounts in United States wheat exports to different countries exist, even after accounting for class differences and the Export Enhancement Program. Further empirical work should investigate if significant country effects reflect quality differences among wheat exports.²

Model Extensions

While Knetter's model proposes a distinction between a competitive and imperfectly competitive market structure, large buyers may behave as monopsonists. Carter and Schmitz, and Love and Murniningtyas have suggested such behavior in the international wheat market. An im-

² See Veeman and Wilson for the effects of different quality characteristics on the international wheat prices.

porter's influence on its purchase prices may depend on its relative market share.

To examine the impact of monopsony power in international wheat trade, a term which measures the import share of each country from the United States (r_{it}) was added to equation (5). This variable was added to capture the possibility that large importers are able to capture monopsony rent. Thus, if a country can exercise monopsony power and obtain a lower price, the coefficient on the import share would be negative and significant.

Table 6 presents the estimated share coefficients in the wheat market using both real and nominal exchange rate measures in equation (5). The share coefficients for the People's Republic of China were negative and significant in both regressions. The Soviet Union's share coefficient was negative and significant when using the real exchange rate measure. Because both China and the Soviet Union are the largest importers of wheat, these results suggest some exercise of monopsony power. The estimates for Japan did not yield a significant share coefficient. Thus, the conclusions by Carter and Schmitz, and Love and Murniningtyas, which suggested that Japan exerts monopsony power in the international wheat market, were not supported by this analysis.

Additional support for these findings was provided by Blandford, who estimated market share model for wheat and corn. Based on the results, the importing countries are ranked according to

criteria which reflect favorable conditions for the United States in each of the markets. The wheat export markets of the Soviet Union and the PRC ranked almost last (exceeding only the EEC wheat market) compared to the other markets considered.

Conclusions

The competitive structure of agricultural exports from the United States is examined using a model of exporter behavior based on pricing decisions across destination markets. Market power in international agricultural trade is revealed in the adjustment patterns of export prices in response to currency movements. A pooled cross-section-time-series regression model is specified and econometric tests are presented to distinguish between a competitive market and two models of imperfect competition.

The results reject the hypothesis that the export pricing decisions by U.S. firms are consistent with price discrimination across destination markets for cotton, corn, and soybeans. The results are ambiguous for the soybean oil, cake, and meal markets, indicating potential price discrimination against Canada and West Germany.

The strongest evidence against the competitive market structure is obtained for international trade in wheat. A share variable accounts for the impact of large importers of wheat from the United States. The share coefficients are negative and significant for the PRC and the Soviet Union, the two largest importers. As their import shares from the United States increase, these countries obtain lower prices for their imports.

A reviewer has suggested that price discrimination may occur in the margins between CIF and FOB prices. Large multinational exporting firms dominate the market for chartering and shipping services and may exercise market power by influencing the margins. However, Binkley and Harrer noted that the structure of the shipping industry ensures that freight rates remain flexible and do not deviate significantly from the costs of shipping. This lends support for the model of pricing to market developed here based on prices which are exclusive of loading or transportation costs. However, an important research topic is to formally develop tests for price discrimination based on the margins between CIF and FOB prices.

This research highlights the link between industrial organization models and export-pricing

Table 6. Market Share Coefficients in the Wheat Market

Destination	Nominal Exchange Rate	Real Exchange Rate
China	-0.560* (-2.431)	-0.454* (-2.148)
Egypt	-0.225 (-0.458)	-0.673 (-1.314)
Japan	-0.334 (-0.665)	-0.443 (-0.809)
South Korea	-0.229 (-0.266)	-0.629 (-0.708)
Philippines	0.220 (0.135)	0.305 (0.174)
USSR	-0.240 (-1.325)	-0.358* (-1.962)
Taiwan	2.360 (1.461)	2.060 (1.198)
Venezuela	0.115 (0.063)	-1.303 (-0.665)
	$R^2 = 0.90$	$R^2 = 0.90$

Note: Values in parentheses are t -values. Asterisks denote t -values significant at the 10% level.

decisions in international agricultural trade. Future work might examine the relationship between exchange rate adjustments and pricing to market in specific industries, focusing on market concentration within an industry and alternative sources of supply. The extent of pricing to market across different commodities and in different countries should also be investigated.

[Received December 1989; final revision received May 1990.]

References

- Alaouze, C. M., A. S. Watson, and N. H. Sturgess, "Oligopoly Pricing in the World Wheat Market." *Amer. J. of Agr. Econ.* 60(1978):173-85.
- Binkley, J. K., and B. Harrer. "Major Determinants of Ocean Freight Rates for Grains: An Econometric Analysis." *Amer. J. Agr. Econ.* 63(1981):47-58.
- Blandford, D. "Market Share Models and the Elasticity of Demand for U.S. Agricultural Exports." *Elasticities in International Agricultural Trade*, ed. C. A. Carter and W. H. Gardiner. Boulder CO: Westview Press, 1988.
- Carter, C., and A. Schmitz. "Import Tariffs and Price Formation in the World Wheat Market." *Amer. J. Agr. Econ.* 61(1979):517-22.
- Dornbusch, R. "Exchange Rates and Prices." *Amer. Econ. Rev.* 77(1987):93-106.
- Farnsworth, H. C. "International Wheat Agreements and Problems." *Quart. J. Econ.* 66(1956):217-48.
- Feinberg, R. M. "The Interaction of Foreign Exchange and Market Power Effects on German Domestic Prices." *J. Indust. Econ.* 35(1986):61-70.
- Karp, L. S., and A. F. McCalla. "Dynamic Games and International Trade: An Application to the World Corn Market." *Amer. J. Agr. Econ.* 65(1983):541-56.
- Knetter, M. M. "Price Discrimination by U.S. and German Exporters." *Amer. Econ. Rev.* 79(1989):198-210.
- Kolstad, C. D., and A. E. Burris. "Imperfectly Competitive Equilibria in International Commodity Markets." *Amer. J. Agr. Econ.* 68(1986):25-36.
- Krugman, P. "Pricing to Market When the Exchange Rate Changes." *Real-Financial Linkages Among Open Economies*, ed. S. W. Arndt and J. D. Richardson. Cambridge MA: MIT Press, 1987.
- Love, A. H., and E. Murniningtyas. "Japanese Wheat Imports: A Test of Market Power." Paper presented at the AAEA annual meeting, Baton Rouge LA, 30 July-2 Aug. 1989.
- McCalla, A. F. "A Duopoly Model of World Wheat Pricing." *J. Farm Econ.* 48(1966):711-27.
- Mendelsohn, C. "Approaches to International Trade Under Nonprice Competition." *J. Farm Econ.* 39(1957):1724-31.
- Mitchell, D. O., and R. C. Duncan. "Market Behavior of Grains Exporters." *World Bank Res. Observer* 2(1987):3-21.
- Paarlberg, P. L. and P. C. Abbott. "Oligopolistic Behavior by Public Agencies in International Trade: The World Wheat Market." *Amer. J. Agr. Econ.* 68(1986):528-42.
- Pick, D. H., and D. Skully. "The Price Structure of the U.S. Wheat Export Market." Washington DC: U.S. Department of Agriculture, Econ. Res. Serv., 1989.
- Sarris, A. H., and J. Freebairn. "Endogenous Price Policies and International Wheat Prices." *Amer. J. Agr. Econ.* 65(1983):214-24.
- Seitzinger, A. H., and P. L. Paarlberg, "A Simulation Model of the U.S. Export Enhancement Program for Wheat." *Amer. J. Agr. Econ.* 72(1990):95-103.
- Veeman, M. M. "Hedonic Price Functions for Wheat in the World Market: Implications for Canadian Wheat Export Strategy." *Can. J. Agr. Econ.* 35(1987):535-52.
- Wilson, W. W. "Differentiation and Implicit Prices in Export Wheat Markets." *West. J. Agr. Econ.* 14(1989):67-77.

Are More Exports Always Better? Comparing Cash and In-Kind Export Subsidies

Robert G. Chambers and Philip L. Paarlberg

Cash and in-kind export subsidies are compared. It is found that the ability to enhance real farm income via export subsidies is sensitive both to the form (cash vs. in-kind) and the point of policy intervention. In-kind subsidies of the type granted under the Export Enhancement Program can in fact reduce real-farm income. Export subsidies are examined under three separate scenarios: no target price or nonrecourse loan programs, a target price but no nonrecourse loan program; and a nonrecourse loan program but no target price program. The presence of a nonrecourse loan program dampens the impacts of both cash and in-kind export subsidies.

Key words: agricultural trade, export subsidies, Export Enhancement Program.

Before the introduction of the Export Enhancement Program (EEP), two types of export subsidies had been used by the United States: a direct cash subsidy for wheat exports and an export payment-in-kind (PIK) for feed grains, cotton, rice, and nonfat dry milk. Under the original export PIK program, exporters showing proof of export sales were given certificates redeemable for Commodity Credit Corporation (CCC) commodities to be exported (Cochrane and Ryan). The EEP reintroduced in-kind export subsidies in 1985.

This paper compares the relative effectiveness of cash and PIK export subsidies in increasing agricultural exports, real prices, and real incomes. Because agricultural policy ultimately must be judged by how it affects agriculture's relative position in the economy, the framework is general equilibrium. The model considers the interplay between farmers, international middlemen, nonfarm producers, domestic consumers, and foreign consumers. The analysis shows that the ability to enhance real farm income is sensitive both to the form (cash subsidies versus PIK subsidies) and point of policy intervention

(at the farm or export level). In particular, the results demonstrate that PIK subsidies of the type granted under the EEP can decrease real farm income. From production agriculture's perspective, an export PIK program (even though it raises total agricultural exports) is not a sure path to improving agriculture's position in the general economy. Sometimes less exports are better.

We first introduce the model. The second section reexamines, solely for comparative purposes, the consequences of introducing cash subsidies to commercial exporters. The third section examines in-kind export subsidies given to farmers, and the fourth considers in-kind subsidies to commercial traders. The fifth section reexamines these results when the real target price of the agricultural good is fixed. Next, tied PIK export subsidies are considered. Finally, the effects of contemporaneous existence of price-support operations on the results are discussed.

The Model

Technology and Preferences

There are two countries. In the home country, four goods are produced: a nontraded agricultural intermediate good, a nontraded nonagricultural intermediate good, a traded agricultural good, and a traded nonagricultural good. Intermediate-goods production is governed by the input-nonjoint producible output set:

Robert G. Chambers is a professor of agricultural and resource economics, University of Maryland; Philip L. Paarlberg is an associate professor, Department of Agricultural Economics, Purdue University.

Scientific Article No. A6104, Contribution No. 8269 of the Maryland Agricultural Experiment Station.

The authors would like to thank, without implicating them, Bruce Gardner and Brian Fisher for helpful comments.

$$(1) \quad Y(L, K_1, K_2) = \left\{ (z_1, z_2) : z_i \in Y_i(L_i, K_i) \ (i = 1, 2); \sum_{i=1}^2 L_i \leq L \right\},$$

where $z_i \in \mathbb{R}_+$, $L_i \in \mathbb{R}_+$, $K_i \in \mathbb{R}_+$ ($i = 1, 2$) and $Y_i(K_i, L_i)$ is a scalar-output producible output set satisfying usual properties (Chambers). The index 1 refers to the intermediate agricultural good. L_i and K_i are a mobile factor devoted to the production of good i and the total endowment of a nontraded i th sector-specific factor, respectively. The total endowment of the mobile factor of production, L , is in fixed supply and not traded internationally. The production technology is thus of the Ricardo-Viner type familiar from trade theory (Jones, Dixit and Norman). Moreover, assume that

$$(2) \quad Y_i(tK_i, tL_i) = tY_i(K_i, L_i) \quad t > 0.$$

Production of each intermediate output satisfies constant returns to scale.

A dual representation of the technology is the restricted profit function:

$$(3) \quad \begin{aligned} \pi_i(p_i, w; K_i) &= \max_{L_i, z_i} \{p_i z_i - w L_i : y_i \in Y_i(K_i, L_i)\} \\ &= K_i \pi_i(p_i, w), \end{aligned}$$

where w is the price of the mobile factor of production and p_i ($i = 1, 2$) is the price of the i th intermediate good. The linearity of (3) in K_i follows from constant returns to scale. Expression (3) is the quasi-rent accruing to the i th sector-specific factor and $\pi_i(p_i, w)$ ($i = 1, 2$) is the rental rate of the i th sector-specific factor of production. Each $\pi_i(p_i, w)$ is convex and positively linearly homogenous in its arguments, nondecreasing in p_i , and nonincreasing in w . Moreover, assuming that a unique solution to (3) exists implies by Hotelling's lemma that (3) is differentiable and that

$$(4) \quad L_i(p_i, w, K_i) = -K_i \frac{\partial \pi_i(p_i, w)}{\partial w} \quad i = 1, 2,$$

where $L_i(p_i, w, K_i)$ is the profit-maximizing derived demand for the mobile factor in the i th intermediate industry.

Because $K_i \pi_i(p_i, w)$ is the quasi-rent accruing to the owners of the i th sector-specific input, it is interpreted as the income accruing to the i th sector. Thus, $K_1 \pi_1(p_1, w)$ is farm or agricultural income. (For example, K_1 can be interpreted as

agricultural land.) Using the Ricardo-Viner technology permits identification of how different sectors' real incomes are affected by export policies. Thus, the focus is on the real-income distributional effects of these programs. No welfare analysis based on a community indifference curve is presented for two reasons: (a) because all the policies considered below amount to the domestic country giving away real income [see equations (9)], the competitive assumptions of the model insure that the country must fall to a lower indifference curve as a result of the policy; (b) the aggregation conditions required for the existence of such a community indifference curve [individual expenditure functions must be of the GL form (Muellbauer)] are so restrictive that little realism is gained by their imposition.

Domestic demand for the i th intermediate good is given by the well-behaved demand functions $c_i(p_1, p_2, E)$ ($i = 1, 2$), where E is expenditure on goods 1 and 2. Each c_i is homogenous of degree zero in its arguments and is integrable. Because the first intermediate good is an agricultural good, Engel's law suggests that its income elasticity is small. Coupling this low elasticity with the relatively small percentage of income spent by U.S. consumers on food suggests that little generally is lost by assuming that $\partial c_1 / \partial E = 0$.¹

National income (I) is the sum of payments to the mobile factor of production and quasi-rents to the sector-specific factors of production:

$$(5) \quad I \equiv \sum_{i=1}^2 K_i \pi_i(p_i, w) + wL.$$

The traded goods are processed versions of the intermediate goods that are ready for export. Let y_a and y_b be the quantities of the traded agricultural and nonagricultural commodities, respectively. For simplicity, the process of converting the intermediate agricultural and nonagricultural products into final exports delivered to the foreign country is represented by the constant-returns-to-scale technologies:

$$(6) \quad \begin{aligned} y_a &= \alpha y_1 \\ y_b &= \beta y_2. \end{aligned}$$

This technology is assumed to include transpor-

¹ This assumption is made mainly to sharpen the analysis and reduce notational clutter. All of the results that follow generally are ambiguous if income effects on the agricultural good are allowed. However, so long as these effects are small, which empirical evidence (George and King) suggests is the case, they will be dominated by the more direct price effects discussed below.

tation and marketing services for trade which otherwise leave the commodity unchanged.²

International demands for commodities a and b are given by the foreign excess demand functions:

$$q_i = q_i(p_a^*/p_b^*) \quad (i = a, b).$$

Here p_i^* represents the price of commodity i in the foreign country. It is assumed throughout that $q_i(\cdot)$ is an excess demand function such that $q'_b > 0$ and $q'_a < 0$. (Again, to avoid complicating the model we have presumed foreign income effects are negligible.)

Equilibrium

All markets are perfectly competitive. This assumption may appear strong given the Schmitz et al. allegation that the international grains market is characterized by quasi-monopolistic traders. The assumption is made for three reasons. It allows isolating the effects of export subsidies without confounding them with the effects of monopolistic practices by international middlemen. Thus, if a policy is found to affect farmers adversely, it can be attributed to the policy and not the actions of the middlemen. If farmers cannot gain in the presence of competitive middlemen, they are even less likely to gain with noncompetitive middlemen. Second, Bieri and Schmitz; McCalla; Alaouze, Watson, and Sturgess; and Chambers and Woolverton have already analyzed the implications of various policy alternatives in the presence of international grain cartels and quasi-monopolistic traders. In particular, Chambers and Woolverton have shown that the institution of a grain cartel could hurt farmers. Third, the grain traders maintain that they are low mark-up pricers making their profits on volume and not price manipulation. Caves presents empirical evidence supporting this claim. This characterization fits the technology described by (6) above.

The equilibrium conditions for the model in the absence of government intervention are the following:

(7)

Mobile factor

$$\sum_{i=1}^2 K_i \frac{\partial \pi_i(p_i, w)}{\partial w} = -L$$

Intermediate goods

$$K_1 \frac{\partial \pi_1(p_1, w)}{\partial p_1} = c_1(p_1, p_2, E) + ay_a$$

$$K_2 \frac{\partial \pi_2(p_2, w)}{\partial p_2} = c_2(p_1, p_2, E) + by_b$$

Traded goods

$$y_a = q_a(p_a^*/p_b^*)$$

$$y_b = q_b(p_a^*/p_b^*)$$

International middlemen

$$p_a y_a - p_1 a y_a = 0$$

$$p_b y_b - p_2 b y_b = 0$$

Budget constraint

$$E = I,$$

where $a = 1/\alpha$ and $b = 1/\beta$.

The first equation in (7) is the equilibrium condition for the mobile factor of production in the home country. The second and third equations state that domestic production of the intermediate goods equals demand by domestic consumers for the intermediate goods plus the derived demand by international middlemen for the intermediate goods. The fourth and fifth equations are the market-clearing conditions for the traded goods sectors. The next two equations are the zero-profit conditions required by free entry and exit of middlemen. The last is the domestic consumer budget (balance of trade) constraint.

The equation system in (7) has ten unknowns:

$$(p_1, p_2, p_a, p_b, p_a^*, p_b^*, w, E, y_a, y_b).$$

Absent government intervention, however, competitive behavior insures that $p_i = p_i^*$ ($i = a, b$).³ There are enough equations to solve for all the unknowns.

To develop the relationship between p_a and p_b and p_1 and p_2 , substitute the final-good produc-

² More generally some would argue strongly that fixed costs associated with processing are large, e.g., grain for export. This could be incorporated into the present model by replacing (6) with a quasi-homothetic technology. Expression (6) can also be easily generalized to include competitively provided inputs for a more general production structure without seriously affecting the analysis.

³ The analysis assumes, unlike the 1956 situation but approximately like the 1985 situation, that the domestic and world prices are the same.

tion functions (6) into the middleman's zero-profit conditions. The price of the intermediate goods 1 and 2 equals the price of the corresponding final good times α and β , respectively: middlemen are mark-up pricing. Use the equilibrium conditions in the traded goods market to incorporate the foreign excess demands into the equilibrium conditions for the intermediate goods. Income can be substituted for E in the demand functions by the budget constraint. Thus, the system of equations is reduced to a three-equation system plus the international price linkages:

$$(8) \quad K_1 \frac{\partial \pi_1(\alpha p_a, w)}{\partial w} + K_2 \frac{\partial \pi_2(\beta p_b, w)}{\partial w} = -L$$

$$K_1 \frac{\partial \pi_1(\alpha p_a, w)}{\partial p_1} - c_1(\alpha p_a, \beta p_b, I) - a q_a(p_a^*/p_b^*) = 0$$

$$K_2 \frac{\partial \pi_2(\beta p_b, w)}{\partial p_2} - c_2(\alpha p_a, \beta p_b, I) - b q_b(p_a^*/p_b^*) = 0.$$

The homogeneity properties of demands and supplies imply one price can be selected as the *numéraire*. Let w be the *numéraire* and real-prices $\bar{p}_i = p_i/w$, $\bar{p}_i^* = p_i^*/w$ ($i = a, b$), and $\bar{I} = I/w$. All purchasing power and prices, therefore, are measured in terms of units of L . With little loss of generality, all normalized prices are assumed initially to equal one.

Cash Export Subsidies

Although the general equilibrium effects of export subsidies are well understood, they are briefly reviewed here to permit easy comparison with the PIK subsidy results.

Introducing cash export subsidies for the traded agricultural commodity requires modifying the price-linkage between the domestic and foreign markets. The real per unit subsidy is denoted \bar{s} (we presume that initially $\bar{s} = 0$). The relationship between domestic and international prices is then

$$\bar{p}_a^* = \bar{p}_a - \bar{s}.$$

A cash export subsidy must be financed out of domestic income. Equation (5) is modified to read in normalized terms:

$$(9) \quad \bar{I} = K_1 \pi_1(\alpha \bar{p}_a, 1) + K_2 \pi_2(\beta \bar{p}_b, 1) + L - \bar{s} y_a.$$

The first three terms give the real income earned by factors of production, and the last term is the real export subsidy cost.

The effects of introducing this subsidy are captured by

$$(10) \quad \frac{\partial \bar{p}_a}{\partial \bar{s}} = \frac{1}{\Delta} [K_2 \partial_{wp} \pi_2 \beta (a q_a')] \\ \frac{\partial \bar{p}_b}{\partial \bar{s}} = -\frac{1}{\Delta} [K_1 \partial_{wp} \pi_1 \alpha (a q_a')],$$

where

$$\Delta = K_1 \partial_{wp} \pi_1 \alpha (a q_a' - c_{12} \beta) - K_2 \partial_{wp} \pi_2 \beta (K_1 \partial_{pp} \pi_1 \alpha - c_{11} \alpha - a q_a').$$

Here the notation $\partial_{ij} \pi_k$ means the second partial derivative of π_k with respect to price i and price j .

The profit function $\pi_i(\cdot)$ is convex in p and w . Hence $\partial_{pp} \pi_i$ is nonnegative. The homogeneity and convexity of π_i imply $\partial_{wp} \pi_i < 0$. With only two consumption goods complementarity is excluded. The terms c_{ij} ($j = 1, 2$) are thus the Hicksian-demand slopes (recall $\partial c_i / \partial E = 0$) and hence are of opposite signs with $c_{11} < 0$ and $c_{12} > 0$. Thus, Δ is unambiguously positive.

Inspection of the numerators in (10) now yields unambiguous results: $\partial \bar{p}_a / \partial \bar{s} > 0$ and $\partial \bar{p}_b / \partial \bar{s} < 0$. The real domestic price of the traded agricultural good rises while \bar{p}_b falls. (Qualitative results are summarized in table 1.) Consequently, the price of the agricultural intermediate good rises, and the price of the intermediate nonagricultural good falls.

World market prices react differently than domestic prices. The export subsidy lowers the border price of the agricultural good facing the foreign country. [Note the first expression in (10) is less than one.] The foreign country's border price for the nonagricultural good also falls as \bar{p}_b falls because there is no policy intervention in this market. Consequently, trade expands because the excess demands of the foreign country have negative slopes.

The effect of the cash export subsidy on real sectoral income can now be easily inferred. Because \bar{p}_a and \bar{p}_1 both rise, the properties of π_i imply that real farm income rises with the cash export subsidy. (Recall total real income to the i th specific factor is $K_i \pi_i(\bar{p}_i, 1)$ which is non-decreasing in \bar{p}_i .) On the other hand, because \bar{p}_2 falls real nonagricultural income falls. The real income of the mobile factor, L , falls relative to farm income and rises relative to nonagricultural income. Farmers benefit from a cash export

Table 1. Effects of Alternative Subsidies

	Without Target Price or Loan Rate						With Target Prices					
	\bar{p}_a	\bar{p}_b	\bar{p}_1	\bar{p}_2	$\pi_1(\bar{p}_1, 1)$	Total Real Farm Income	\bar{p}_a	\bar{p}_b	\bar{p}_1	\bar{p}_2	$\pi_1(\bar{p}_1, 1)$	Total Real Farm Income
Cash export subsidy	+	-	+	-	+	+	+	0	+	0	0	0
PIK to farmers	-	+	-	+	-	?	-	0	-	0	0	?
PIK to middlemen	-	?	?	?	?	?	-	0	?	0	0	?
With Loan Rates												
Cash export subsidy	+	-	+	-	+	+						
PIK to farmers	-	+	-	+	-	?						
PIK to middlemen	-	?	?	?	?	?						

subsidy given to the international middleman, while nonfarmers lose.

In-Kind Export Subsidy to Farmers

In modeling in-kind subsidies to farmers, the in-kind subsidy amount is not considered a portion of current supply or demand. It only consists of accumulated government stocks. As was true when the EEP was initiated, these stocks are initially assumed isolated from the market and not price sensitive. Moreover, we assume initially that these stocks do not comprise a portion of current flow income of the government. (Both assumptions are relaxed below.) Thus, the in-kind subsidy yields a rightward shift in intermediate agricultural-commodity supply. The analysis, reminiscent of trade growth models, therefore focuses solely on the consequences of an in-kind export subsidy to which farmers have first title and not a *PIK* program where *PIK* payments compensate producers for retiring acreage.

Denoting the in-kind subsidy as *PIK*, the intermediate agricultural market-clearing condition is respecified

$$(11) \quad PIK + K_1 \frac{\partial \pi_1(\bar{p}_1, 1)}{\partial p} = c_1(\bar{p}_1, \bar{p}_2, \bar{I}) + \alpha y_a.$$

Introducing a *PIK* subsidy is analyzed by differentiating the new equilibrium conditions and evaluating all derivatives at *PIK* = 0 to get

$$(12) \quad \frac{\partial \bar{p}_a}{\partial PIK} = \frac{K_2}{\Delta} \partial_{wp} \pi_2 \beta < 0;$$

$$\frac{\partial \bar{p}_b}{\partial PIK} = \frac{-K_1}{\Delta} \partial_{wp} \pi_1 \alpha > 0.$$

The real price of the traded agricultural commodity falls while the traded nonagricultural commodity real price rises. Contrast this with the cash export subsidy result where \bar{p}_a rises and \bar{p}_b falls. The cash and *PIK* export subsidies have opposite effects. To see why notice that the *PIK*-induced rightward shift in agricultural supply makes the agricultural price fall to soak up the resulting excess supply at original prices. The real nonagricultural price rises to divert demand from the nonagricultural commodity to the agricultural commodity. As a result of these price adjustments, \bar{p}_1 falls while \bar{p}_2 rises. These latter real prices movements imply that the real rental rate on the agricultural-specific factor ($\pi_1(\bar{p}_1, 1)$) falls while $\pi_2(\bar{p}_2, 1)$ rises. Agriculture's share in earned national income falls relative to that of the mobile factor and to the owners of the nonagricultural-specific factor. Again, contrast this with the results for a cash subsidy.

As the real prices of the agricultural traded and intermediate goods fall, so do commercial (nongovernment stock) exports of the traded agricultural commodity. Total agricultural exports rise because of the lower world price of the traded agricultural good relative to the nonagricultural good. Noncommercial (*PIK*) exports actually displace commercial agricultural exports.

Although $\pi_1(\bar{p}_1, 1)$ falls, farmers may not lose because the *PIK* subsidy is made directly to them. Because farmers have title to the *PIK* amount, a necessary condition for their real total income to rise as a result of introducing a *PIK* payment is

$$(13) \quad \frac{\partial(\bar{y}_1)}{\partial PIK} \cdot \frac{PIK}{(\bar{p}_1)} > \frac{-PIK}{y_1}.$$

Expression (13) illustrates that when total demand for the intermediate agricultural commodity is very price inelastic, introducing a *PIK*

subsidy to farmers causes a large fall in \bar{p}_1 . The induced effect on farm income is both large and negative. For farmers to gain from the subsidy, the *PIK* amount must be large relative to farm production. In contrast, when total demand for the intermediate agricultural commodity is very elastic, the *PIK* subsidy acts as a virtual lump-sum transfer to farmers and tends to raise their incomes.

That a *PIK* export subsidy to farmers can make farmers worse off apparently has not been clearly understood. Granting the *PIK* export subsidies to farmers has much the same effect as exogenous growth in a traditional trade model.⁴ There, growth, via an increased endowment or technical change, causes an adverse terms-of-trade shift which counters the gain from the output expansion (Dixit and Norman). If the worsening terms-of-trade effect is large enough to dominate the gain from increased output, national welfare falls—growth is “immiserizing.” Whether growth is immiserizing is governed by the elasticities of the model. Here, however, the *PIK* subsidy always deteriorates the terms of trade which in turn always adversely affects the export sector (agriculture). On balance, this adverse trade effect outweighs the *PIK* giveaway if (13) is not satisfied.

In-Kind Subsidy to Exporters

Now consider what happens when the *PIK* subsidy is paid directly to international middlemen. Three recent analyses consider this type of subsidy. Houck examines a general export bonus scheme where the in-kind payment is tied to commercial exports (also see below). Hillberg, and Seitzinger and Paarlberg model EEP in a bargaining theoretic framework. None of these studies consider the general equilibrium consequences of such *PIK* subsidies.

When the EEP was announced in spring 1985, great care was taken to prevent domestic resale of the *PIK* subsidies. Sales were targeted at specific countries, and in-kind subsidies were granted only to companies which had already made sales agreements with importing nations. The attempt to shield domestic markets from a *PIK* export subsidy was not new. Starting in 1956, a *PIK* payment to subsidize wheat and flour exports had been available to exporters from CCC stocks. Certificates redeemable for CCC stocks were made available to exporters upon proof of ex-

ports from private wheat stocks, but the wheat taken from CCC stocks could only be sold for export (Cochrane and Ryan).

The EEP provisions and the 1956 *PIK* program suggest that whether a *PIK* grant to exporters can be resold domestically has important consequences. But it is easy to show that the fungible nature of internationally traded commodities makes such export provisions irrelevant: the same equilibrium conditions emerge whether exporters can resell the *PIK* grant domestically or not (Chambers and Paarlberg). Here, we only consider a *PIK* grant which must be exported.

Consider a *PIK* export subsidy given to the exporting companies which must be exported. The equilibrium condition for the intermediate agricultural good remains the same, but y_1 must now be interpreted as purchases by the trading firms for “commercial” exports. The equilibrium condition for the traded agricultural good must be modified to reflect the *PIK* subsidy. Total exports now equal “commercial” exports plus *PIK* exports:

$$(14) \quad q_a(p_a^*/p_b^*) = \alpha(PIK + y_1).$$

Because the exporting companies sell the *PIK* but only buy y_1 on the domestic market, their zero profit condition becomes

$$(15) \quad p_a\alpha(PIK + y_1) - p_1y_1 = 0.$$

Solving (14) and (15) recursively gives

$$(16) \quad y_1 = aq_a - PIK,$$

$$(17) \quad p_1 = \frac{\alpha p_a q_a}{q_a - \alpha PIK}.$$

Substituting (16) into the intermediate agricultural good's market clearing condition gives the following equilibrium conditions:

$$(18) \quad K_1 \frac{\partial \pi_1}{\partial p_1} = c_1 + aq_a - PIK,$$

$$\sum_{i=1}^2 K_i \frac{\partial \pi_i}{\partial w} = -L,$$

$$p_1 = \frac{\alpha p_a q_a}{q_a - \alpha PIK}.$$

Normalizing all prices by w and choosing units such that initially $p_a/w = p_b/w = a = 1$ and

⁴ We owe this interpretation to an anonymous reviewer.

$PIK = 0$ gives the following comparative static expression:

$$(19) \quad \begin{bmatrix} 1 & -1 & 0 \\ K_1 \partial_{wp} \pi_1 & 0 & K_2 \partial_{wp} \pi_2 \beta \\ K_1 \partial_{pp} \pi_1 - c_{11} & -q'_a & q'_a - c_{12} \beta \end{bmatrix} \cdot \begin{bmatrix} d(\bar{p}_1) \\ d(\bar{p}_a) \\ d(\bar{p}_b) \end{bmatrix} = \begin{bmatrix} 1 \\ q_a \\ 0 \\ 1 \end{bmatrix} dPIK.$$

The first row of (19) [and the last in (18)] shows that the simple linear link between \bar{p}_1 and \bar{p}_a is now broken. The effect on \bar{p}_1 of a PIK subsidy given to middlemen has two parts:

$$(20) \quad \frac{d(\bar{p}_1)}{dPIK} = \frac{d(\bar{p}_a)}{dPIK} + \frac{1}{q_a}.$$

The first part, through \bar{p}_a , emerged in the previous subsidy cases: the PIK subsidy to farmers and the cash export subsidy. To see the origin of the second term recall the following: an in-kind subsidy to exporters enables them to take a smaller return on each commercial sale than before and still break even. While \bar{p}_a is expected to fall as a result of the subsidy (indeed, this is shown later), the secondary effect can ameliorate, and under suitable conditions even overcome, the first effect.

Solving (19) gives (at $\alpha = 1$):

$$(21) \quad \frac{\partial(\bar{p}_a)}{\partial PIK} = - \frac{1}{q_a} \frac{K_2 \partial_{wp} \pi_2 \beta [(c_{11} - K_1 \partial_{pp} \pi_1) - q_a] + K_1 \partial_{wp} \pi_1 (q'_a - c_{12} \beta)}{K_2 \partial_{wp} \pi_2 \beta [(c_{11} - K_1 \partial_{pp} \pi_1) + q'_a] + K_1 \partial_{wp} \pi_1 (q'_a - c_{12} \beta)},$$

$$(22) \quad \frac{\partial(\bar{p}_b)}{\partial PIK} = \frac{1}{q_a} \frac{K_1 \partial_{wp} \pi_1 [(c_{11} - K_1 \partial_{pp} \pi_1) - q_a] + K_1 \partial_{wp} \pi_1 [K_1 \partial_{pp} \pi_1 - c_{11} - q'_a]}{K_2 \partial_{wp} \pi_2 \beta [(c_{11} - K_1 \partial_{pp} \pi_1) + q'_a] + K_1 \partial_{wp} \pi_1 (q'_a - c_{12} \beta)} \\ = - \frac{1}{q_a} \frac{K_1 \partial_{wp} \pi_1 [q_a + q'_a]}{K_2 \partial_{wp} \pi_2 \beta [(c_{11} - K_1 \partial_{pp} \pi_1) + q'_a] + K_1 \partial_{wp} \pi_1 (q'_a - c_{12} \beta)}.$$

The sign of (21) is always negative. A PIK subsidy to middlemen also has the opposite effect of a cash subsidy on \bar{p}_a . The sign of (22) is positive if the foreign excess demand for the agricultural traded good is inelastic. This is true because the denominator of (22) is positive by earlier arguments, while the numerator's sign depends upon the expression $-[q_a + q'_a]$. Because all prices are normalized to unity initially, this expression is $-q_a[1 + \eta]$, where η is the elasticity of excess demand (q'_a/q_a). If excess

demand is inelastic, (22) is positive (recall $K_1 \partial_{wp} \pi_1$ is negative). (This assumes $q_a > 0$, the home country is a net exporter of a .)

Intuitively, an inelastic agricultural excess demand implies that the PIK -induced supply shift yields little quantity adjustment but relatively large declines in \bar{p}_a . This result translates into reduced foreign expenditure for agricultural products which means greater expenditures on nonagricultural products and higher prices for b . If \bar{p}_b rises, i.e., $-(1 + \eta) < 0$ quasi-rent to the nonagricultural-specific factor rises (p_2/w remains linearly related to p_b/w).

A condition for \bar{p}_1 to rise is

$$\frac{1}{q_a} \left(q_a \frac{d(\bar{p}_a)}{dPIK} + 1 \right) > 0.$$

From above $d(\bar{p}_a) < 0$. A necessary and sufficient condition for (\bar{p}_1) to rise is that the numerator of the right-hand side of (21) be less than the denominator. But the only differences between the terms is the presence of $(-q_a)$ in the numerator and q'_a in the denominator. Hence, (\bar{p}_1) rises only if excess demand for the traded agricultural commodity is elastic. This is similar to Houck's result obtained when p_1 is a blend of the commercial and subsidized prices. If excess demand for the agricultural traded good is inelastic, a PIK subsidy to middlemen causes \bar{p}_1 , and thus, $\pi_1(\bar{p}_1, 1)$ to fall. Real farm income

falls not rises. Farmers are hurt not helped by the PIK subsidy.

These results permit determining the effects of introducing a PIK export subsidy on commercial exports. Commercial exports of the traded agricultural commodity are

$$(23) \quad \text{Commercial exports} = q_a(p_a^*/p_b^*) - PIK.$$

Expression (23) implies the following necessary

and sufficient condition for introducing a *PIK* subsidy to exporters to raise commercial exports:

$$(24) \quad q'_a \left(\frac{d(\bar{p}_a)}{dPIK} - \frac{d(\bar{p}_b)}{dPIK} \right) > 1.$$

Substituting from equations (21) and (22) into (23) yields after a slight rearrangement of terms:

$$(25) \quad \frac{-q'_a \left[K_2 \partial_{wp} \pi_2 \beta (c_{11} - K_1 \partial_{pp} \pi_1 - q_a) + K_1 \partial_{wp} \pi_1 (-c_{12} \beta - q'_a) \right]}{q_a \left[K_2 \partial_{wp} \pi_2 \beta (c_{11} - K_1 \partial_{pp} \pi_1 + q'_a) + K_1 \partial_{wp} \pi_1 (q'_a - c_{12} \beta) \right]} > 1.$$

Because the denominator of (25) is positive, the condition for commercial exports to rise is

$$(26) \quad -(1 + \eta) [K_2 \partial_{wp} \pi_2 \beta (c_{11} - K_1 \partial_{pp} \pi_1) - K_1 \partial_{wp} \pi_1 c_{12}] > 0.$$

The term in square brackets on the left-hand side of (26) is positive. Thus commercial exports fall if excess demand is inelastic, $-(1 + \eta) > 0$.

The assumption of zero trader profit implied by pure competition is important to the results in this section. Suppose, to the contrary, that traders only need a zero profit on their commercial sales and are allowed to retain the value of the *PIK* amount as a lump-sum transfer. In this case, the equilibrium conditions are the same as for an in-kind subsidy to farmers. But now farmers always lose because they see falling prices and incomes and derive no income from the *PIK* amount. This intuition suggests that the less competitive is the international grains market, the more likely it is that a *PIK* subsidy will hurt farmers.

Export Subsidies and a Target Price

The results presented above examine the impact of cash and in-kind export subsidies without considering other policies. But trade policies usually operate in concert with other interventions. In the United States, agricultural producer prices are often guaranteed by setting high target prices which become the effective producer price. In this context, an export subsidy is often the trade manifestation of the domestic intervention. Hillberg, for example, shows that the levels of U.S. domestic policy instruments are critical in empirically evaluating EEP.

To examine the sensitivity of results to domestic policy, the subsidy policies are reanalyzed in the presence of a real target price set

above the market-clearing equilibrium price and maintained by deficiency payments. Price-support operations through a nonrecourse loan program are considered separately below.

In this section the effective producer price for the intermediate agricultural good is the real target price which is policy determined and is unchanged by the subsidy program.⁵ If \bar{p}_1 is above the real target price, the analysis proceeds as be-

fore. Only cases where the target price is effective are of interest. The real target price is assumed fixed. If only the nominal price is fixed (the more usual case), then the real producer price of the intermediate good would be free to vary as in earlier sections.

Domestic real income available for consumption is income less the deficiency payment cost:

$$\bar{I} = \sum_{i=1}^2 K_i \frac{\partial \pi_i(\bar{p}_i, 1)}{\partial w} + L - (\bar{p}_1^s - \bar{p}_1) y_1,$$

where \bar{p}_1^s is the real target price. Differentiating the mobile-factor market equilibrium gives

$$(27) \quad K_1 \partial_{wp} \pi_1 d\bar{p}_1^s + K_2 \partial_{wp} \pi_2 \beta d\bar{p}_b = 0.$$

With the real target price fixed, $d\bar{p}_1^s = 0$, and hence from (27) $d\bar{p}_b = 0$. An export subsidy, therefore, causes no change in domestic relative producer prices. Hence, the export subsidy no longer affects the allocation of L . And, because endowments of K_1 and K_2 are fixed, the export subsidy has no effect on domestic production patterns. This result implies, in turn, that the price of the nonagricultural traded good and its intermediate counterpart cannot change. Therefore, with a fixed real target price, all adjustment comes in terms of the domestic consumption price and the price to foreign buyers of the agricultural intermediate and its traded counterpart.

A cash export subsidy's impact on the tradeable agricultural good's real price to consumers and foreigners is

$$\frac{d\bar{p}_a}{d\bar{s}} = \frac{aq'_a}{c_{11}\alpha + aq'_a}.$$

The denominator is clearly negative because it contains own-price effects from domestic and

⁵ Implicitly we assume all producers are in the program.

excess demand. The numerator is also negative. Thus, \bar{p}_a rises in response to the cash subsidy. As above, however, $d\bar{p}_a/ds < 1$, the domestic price of the agricultural commodity rises by less than the subsidy amount and the price abroad falls.

The rise in \bar{p}_a induces an increase in \bar{p}_1 , which reduces domestic demand for the intermediate agricultural good. Because the producer price is fixed at \bar{p}_1^* , no production change occurs. No matter what happens, producers continue to receive the target price. With static production and falling domestic consumption, exports of the agricultural good must rise. (This also can be seen by noting that \bar{p}_a^* falling causes an expansion of trade.) Producer income and the returns to sector-specific factors are the same as before the subsidy policy. A cash export subsidy in the presence of a fixed real target price only acts to increase agricultural exports, it does not improve agriculture's position in the national economy. Subsidy rents are transferred to foreign consumers.

With an in-kind subsidy to farmers, the associated change in \bar{p}_a is

$$(28) \quad \frac{d\bar{p}_a}{dPIK} = (c_{11}\alpha + aq'_a)^{-1}.$$

Expression (28) shows that \bar{p}_a falls in response to the in-kind subsidy. This price decline is necessary to absorb the additional supply of government stocks placed on the market. Total exports are greater as \bar{p}_a^* is also lower.

The consumer price of the agricultural intermediate good also falls. Domestic use of the intermediate rises. Because production is unchanged with the fixed target price and domestic use is greater, "commercial" exports are lower both as a share of total exports and in absolute volume. As before, commercial exports are displaced by exports from released government stocks.

With \bar{p}_1^* constant, sector-specific factor prices remain unchanged by the policy. However, the income of agricultural producers now clearly rises as their volume marketed expands by the amount PIK , which is sold at price \bar{p}_1 . Agriculture, therefore, gains as a result of the export subsidy.

When the in-kind subsidy is given to the international middleman,

$$\frac{d\bar{p}_a}{dPIK} = \frac{\left(-1 + c_{11}\frac{\alpha}{y_1}\right)}{-(c_{11}\alpha + aq'_a)},$$

which is negative. Because \bar{p}_a falls and there is no price wedge in the world market, \bar{p}_a^* falls and total exports rise. With the real target price constant, producers of the intermediate agricultural commodity experience no changes in this policy environment. Subsidy benefits are transferred to domestic and foreign consumers.

Considering the in-kind subsidy to the middleman without the target price, the change in \bar{p}_1 is ambiguous. This result occurs because of the conflict between the negative terms-of-trade effect and the positive direct effect. A similar situation holds here. The change in \bar{p}_1 faced by domestic consumers is

$$\frac{d\bar{p}_1}{dPIK} = \frac{\alpha \left(1 + q'_a \frac{a}{y_1}\right)}{(c_{11}\alpha + aq'_a)}.$$

If excess demand for the tradeable agricultural good is elastic, total exports increase enough to increase the domestic consumption price of the intermediate good. This occurs because the increase in total trade requires an increase in commercial exports which can only be met through a reduction in domestic use. If the excess demand is not elastic, then the change in the domestic price of the intermediate agricultural good depends on the ratio of total trade to commercial trade— $aq'_a(\cdot)/y_1$. Larger in-kind subsidies will raise this ratio by lowering y_a , and hence, will be more likely to raise \bar{p}_1 .

Tied PIK Export Subsidies

So far, all PIK export subsidies have been lump-sum.⁶ While this specification closely approximates the EEP type of PIK subsidy, other forms of PIK subsidies are possible. This section considers PIK subsidies tied to the level of agricultural exports. Only the case where the PIK subsidy is granted to farmers is analyzed. The reader can easily extend the analysis to the case where the PIK subsidy is granted to exporters using previous developments.

⁶ We would like to thank an anonymous reviewer for suggesting the analysis contained in this section.

Let τ denote percent of total exports comprised by the *PIK* subsidy, i.e.,

$$PIK = \tau q_a.$$

This necessitates the following respecification of (11):

$$(11') \quad K_1 \frac{\partial \pi_1(\bar{p}_1, 1)}{\partial p} = c_1(\bar{p}_1, \bar{p}_2, \bar{I}) + (a - \tau)y_a.$$

Differentiating the new equilibrium conditions and evaluating all derivatives at $\tau = 0$ gives the following effects of introducing a *PIK* subsidy tied to total exports:

$$(29) \quad \begin{aligned} \frac{\partial \bar{p}_a}{\partial \tau} &= \frac{q_a K_2 \partial_{wp} \pi_2 \beta}{\Delta}, \\ \frac{\partial \bar{p}_b}{\partial \tau} &= \frac{-q_a K_1 \partial_{wp} \pi_1 \alpha}{\Delta}. \end{aligned}$$

Comparing (29) with (12) reveals that a tied export *PIK* subsidy granted to farmers has the same qualitative effects as a lump-sum export *PIK* subsidy granted to farmers. (This presumes, of course, that when the *PIK* subsidy is granted the subsidizing country exports the agricultural commodity, i.e., $q_a > 0$.)

A Longer-Run View

Thus far, granting the *PIK* subsidy has had no effect on current national income because a *PIK* subsidy is assumed to enter the national income accounts as either transfers from the government to farmers or to exporters.⁷ The effect on national income is a wash unless the demand functions are specified to take income distributional issues into account. A primary justification for this specification is that the original EEP program used stocks accumulated in historical periods by the government. Current flow income was not affected by those transfers. But this analysis ignores an important point. The EEP has not been a one-time export subsidy pro-

gram. Rather it evolved into an ongoing program extending over at least the life of the Food Security Act of 1985.

If all EEP subsidies were granted from government stocks accumulated before the institution of EEP, the analysis would proceed as before. But current EEP subsidies are drawn from government stocks continuously acquired through the government's price-support (typically non-recourse loan) operations. These price-support operations, in turn, are affected by the *PIK*-subsidy program. As *PIK*-subsidies drive market prices down more stocks are acquired through forfeited loans. This has two effects: the price-support operations mitigate the degree to which agricultural prices fall as a result of the *PIK* subsidy and the government incurs additional budgetary expenses out of current flow income to support market prices.

To assess the longer-run consequences of an export *PIK* subsidy, therefore, the government's price-support operations must be incorporated into the analysis. A nonrecourse loan program is introduced in which the government agrees to purchase all commodities offered at a guaranteed price—the loan rate. The CCC regularly acquires stocks of supported agricultural commodities at prices above the loan rate (even after corrected for interest costs, etc.). Therefore, current government acquisitions of stocks are assumed to be governed by $i(p_1/p_l)$, where p_l is the loan rate. We make the following assumptions:

$$i'(k) < 0, \text{ and}$$

$$\lim_{k \rightarrow 1} i'(k) = -\infty.$$

The first assumption means that government accumulation of inventories through forfeited loans declines as p_1 rises relative to the loan rate. The second assumption means that CCC stands ready to buy all of the commodity offered to it at the loan rate: as the market price approaches the loan rate, government demand for inventories becomes perfectly elastic.⁸

These price-support operations are introduced into preceding analyses. The most obvious change is that expression (5) must be rewritten to rec-

⁷ This analysis was suggested to us by an anonymous reviewer.

⁸ Similar specifications are embedded in the FAPRI and Wheat-sim policy models.

ognize the effect loan defaults have on flow income:

$$(5') \quad I = \sum_{i=1}^2 K_i \pi_i(p_i, w) + wL - p_i i(p_i/p_i).$$

The other change is that the intermediate agricultural good equilibrium must be changed to reflect the new source of government demand for the commodity. Hence, rewrite that condition as

$$K_1 \frac{\partial \pi_1(p_1, w)}{\partial p_1} = c_1(p_1, p_2, E) + i(p_1/p_i) + ay_a.$$

The first step is to analyze how the comparative static effects of a cash export subsidy change as a result of the price-support operations. The necessary calculations give the following results:

$$(10') \quad \frac{\partial \bar{p}_a}{\partial \bar{s}} = \frac{K_2 \partial_{wp} \pi_2 \beta (q'_a a)}{\Delta(p_1/p_i)}, \text{ and}$$

$$\frac{\partial \bar{p}_b}{\partial \bar{s}} = \frac{K_1 \partial_{wp} \pi_1 \alpha (q'_a a)}{\Delta(p_1/p_i)}$$

where

$$\Delta(p_1/p_i) = \Delta + \frac{K_2 \partial_{wp} \pi_2 \beta i'(\bar{p}_1/\bar{p}_i) \cdot \alpha}{\bar{p}_i}.$$

By earlier results, $\Delta(p_1/p_i) > \Delta$. And by the assumptions on $i(p_1/p_i)$, it follows that

$$\lim_{k \rightarrow 1} \Delta(k) = \infty.$$

Therefore, as expected, allowing for price-support intervention dampens the price responsiveness to export subsidies in both markets. In the limit, where the market price is "riding" the loan rate, a cash subsidy has no effect on market prices. Because both \bar{p}_a and \bar{p}_b change less than in the absence of price-support intervention, a cash export subsidy is less effective in raising farm income in the presence of a price-support program than in its absence. In contrast, non-farm prices (and nonfarm income) fall less than in the absence of such programs.

Turning to a *PIK* export subsidy given to farmers,

$$(12') \quad \frac{\partial \bar{s}_a}{\partial FIK} = \frac{K_2 \partial_{wp} \pi_2 \beta}{\Delta(p_1/p_i)},$$

$$\frac{\partial \bar{s}_b}{\partial FIK} = \frac{-K_1 \partial_{wp} \pi_1 \alpha}{\Delta(p_1/p_i)}.$$

Again we conclude that price-support programs dampen the price response to the *PIK* program. Real agricultural prices fall less than they would have in the absence of the price-support program and real nonagricultural prices rise less. As the initial p_1 approaches the loan rate, a *PIK* program given to farmers has no price effect in either agricultural or nonagricultural markets. And, in general, because \bar{p}_1 falls less than in the absence of price support, it is more likely, *ceteris paribus*, that a *PIK* subsidy to farmers raises total farm income when a price-support program exists. The price-support operations dampen the adverse terms of trade effect caused by the *PIK* subsidy.

The effect on total government inventories consists of two parts: the *PIK* giveaway and the reaccumulation of inventories through the price-support mechanism. The expression for the total effect is

$$(30) \quad \frac{i'(\bar{p}_1, \bar{p}_i) \alpha K_2 \partial_{wp} \pi_2 \beta}{\bar{p}_i \Delta(p_1/p_i)} - 1.$$

Using (10') and (12') implies that expression (30) is always nonpositive. A *PIK* export subsidy given to farmers always ends with the government not increasing commodity stocks. The long-run implication is that, *ceteris paribus*, an export *PIK* given to farmers eventually dissipates the government's inventories as less and less beginning inventory is available each period to subsidize exports.

As the initial market price approaches the loan rate, the first term in (30) approaches unity. Therefore, if the market price is "riding the loan rate" when the *FIK* giveaway is instituted, all the *PIK* commodity given away is immediately reabsorbed through the government's price-support operations. In this limiting case, the *PIK* giveaway acts as a pure lump-sum transfer program to farmers. Farmers receive a lump-sum transfer equalling $p_i \cdot PIK$. Under these circumstances the demand faced by farmers is perfectly elastic as the CCC buys everything that is offered at p_i . Thus, expression (13) is satisfied.

Finally, consider how a price-support program changes the results for a *PIK* export subsidy granted to exporting companies. Performing the appropriate comparative-static manipulations gives (at $\alpha = 1$):

$$(21') \quad \frac{\partial \bar{p}_a}{\partial PIK} = - \frac{1}{q_a} \frac{[K_2 \partial_{wp} \pi_2 \beta (c_{11} + i'(\bar{p}_1/\bar{p}_l)/\bar{p}_l - K_1 \partial_{pp} \pi_1 - q_a) + K_1 \partial_{wp} \pi_1 (q'_a - c_{12} \beta)]}{\Delta(p_1/\bar{p}_l)},$$

$$(22') \quad \frac{\partial \bar{p}_b}{\partial PIK} = - \frac{1}{q_a} \frac{K_1 \partial_{wp} \pi_1 (q_a + q'_a)}{\Delta(p_1/\bar{p}_l)}.$$

By (22') the price response in the nonagricultural traded goods market is dampened by the introduction of the price-support operations. However, while the traded agricultural good price still falls as a result of the *PIK* export subsidy, it is ambiguous whether it falls by more or less than in the absence of the price-support program. However, in the limit as p_1 approaches the loan rate, \bar{p}_a falls $(-1/q_a)$ units for every unit of the *PIK* subsidy. But as (20) verifies, this last result implies that \bar{p}_1 does not change and, consequently, neither does farm income in this limiting case. Total farm income remains constant. So, in the case which most closely approximates the 1985 situation (market prices riding the loan rate and *PIK* subsidies given to exporters), farmers realize no gain from the *PIK* subsidy. Instead, the export *PIK* subsidy to exporters ends as a lump-sum transfer to domestic and foreign consumers (absent a competitive grains market, it would be a lump-sum transfer to exporters).

Conclusions

This paper demonstrates the differences between cash and in-kind export subsidies. These results are summarized in table 1. Cash export subsidies expand exports by raising domestic prices and lowering the world price. The price of the agricultural product in the subsidizing country rises, thereby benefiting farm producers at the expense of consumers and taxpayers. An in-kind subsidy given to farmers lowers the price received by farmers but expands total farm exports. The *PIK* commodities displace commercial exports. Although real farm production income is lower after the *PIK*, farmers are not necessarily worse off because the in-kind payment is made directly to them. When international excess demand is elastic, farmers may gain from the introduction of *PIK* payments. An in-

kind export subsidy given to competitive international middlemen lowers the world price and expands trade. The effect of the subsidy on farmers is ambiguous. If the excess demand for the agricultural good is elastic, then agricul-

ture's share of national income rises. For an inelastic excess demand, farm income falls.

Analysis of these three subsidy alternatives in the presence of an effective target price leaves most of the results unchanged. However, with the producer price determined by the target price, producer income is not altered by cash export subsidies and in-kind subsidies to international traders. When the in-kind payment is made directly to farmers, producer income is unambiguously higher with an effective target price.

When a price support program is in effect, the impacts of such subsidies are dampened. Government inventories are reduced by in-kind subsidies over the long run. Further, the analysis shows when the market price approaches the loan rate, the *PIK* subsidy given to farmers becomes a lump-sum subsidy, while the cash and in-kind subsidy to traders leave farm income unaffected. The *PIK* commodities given away are immediately reabsorbed through the government's price support operations.

[Received March 1989; final revision received March 1990.]

References

- Alaouze, C., A. Watson, and N. Sturgess. "Oligopoly Pricing in the World Wheat Market." *Amer. J. Agr. Econ.* 60(1978):173-85.
- Bieri, J., and A. Schmitz. "Market Intermediaries and Price Instability: Some Welfare Implications." *Amer. J. Agr. Econ.* 56(1974):280-85.
- Caves, R. E. "Organization, Scale, and Performance of the Grain Trade." *Food Res. Inst. Stud.* 16(1977-78).
- Chambers, R. G. *Applied Production Analysis: A Dual Approach*. New York: Cambridge University Press, 1988.
- Chambers, R. G., and P. L. Paarlberg. "Are More Exports Always Better: Comparing Cash and In-Kind Export Subsidies." Dep. Agr. and Resour. Econ. Work. Pap. No. 89-02, University of Maryland, Jan. 1989.
- Chambers, R. G., and M. W. Woolverton. "Wheat Cartelization and Domestic Markets." *Amer. J. Agr. Econ.* 62(1980):629-38.

- Cochrane, W. W., and M. E. Ryan. *American Farm Policy: 1948-1973*. Minneapolis: University of Minnesota Press, 1976.
- Dixit, A., and V. Norman. *Theory of International Trade*. New York: J. Nisbet/Cambridge University Press, 1980.
- George, P. S., and G. A. King. *Consumer Demand for Food Commodities in the United States with Projections for 1980*. Giannini Foundation Monograph No. 26. Agricultural Experiment Station, University of California.
- Hillberg, A. M. "The United States' Export Enhancement Program for Wheat: A Simulation Model Employing Nash's Bargaining Solution," Ph.D. thesis, Purdue University, 1988.
- Houck, J. P. "The Basic Economics of an Export Bonus Scheme." *N. Cent. J. Agr. Econ.* 8(1986):227-35.
- Jones, R. H. "Distortion in Factor Markets and the General Equilibrium Model of Production," *J. Polit. Econ.* 79(1971):437-59.
- McCalla, A. "A Duopoly Model of World Wheat Pricing." *J. Farm Econ.* 48(1966):711-27.
- Muellerbauer, J. "Community Preferences and the Representative Consumer." *Econometrica* 44(1976):979-99.
- Schmitz, A., A. F. McCalla, D. O. Mitchell, and C. Carter. *Grain Export Cartels*. Cambridge MA: Ballinger Publishing Co., 1981.
- Seitzinger, A. H., and P. L. Paarlberg. "A Simulation Model of the U.S. Export Enhancement Program for Wheat." *Amer. J. Agr. Econ.* 72(1990):95-103.
- Shubik, M., with R. Levitan. *Market Structure and Behavior*. Cambridge MA: Harvard University Press, 1980.

Testing for Input Substitution in a Regulated Fishery

Diane P. Dupont

Input restrictions are commonly used to prevent rent dissipation in fisheries. This paper examines whether these schemes are successful by calculating the degree of input substitution between restricted and unrestricted inputs. Conventional elasticities of substitution cannot be used when the firm faces controls on the use of some inputs. In this case, the appropriate measure of substitution, the elasticity of intensity, must be used. Data from the British Columbia salmon fishery provide evidence of input substitution possibilities for two vessel types. These results call into question the usefulness of input control schemes.

Key words: fisheries, input substitution, normalized quadratic profit function, partial static equilibrium, regulation.

Fish resources can be a valuable source of revenue to cash-strapped governments if the fisheries are regulated efficiently. In the first-best optimum solution, a sole owner maximizes resource rent. However, without property rights to fish, competition among fishermen can result in a complete dissipation of rent (Warming). To preserve resource rents, regulators often restrict inputs (Christy, Karpoff). Typically, control schemes begin as limited entry programs that restrict only the number of vessels allowed to participate in (or enter) the fishery. Subsequently, restrictions upon input usage per vessel are imposed when regulators find that the first set of restrictions does not work (Pearse, Rettig). The ultimate success of input restrictions appears to rest upon whether fishermen can find substitutes for the restricted inputs.

Past research is divided over the substitution possibilities open to fishermen. Crutchfield argues that fishermen face a sharply increasing

marginal cost of expanding against a tonnage restriction which suggests that the fishing vessel's tonnage is used in fixed proportions with all other inputs, e.g., fuel, labor, gear/equipment.¹ On the other hand, Scott claims that the fishing technology allows much scope for substitution. Until recently, empirical evidence has not been available to decide the issue (Strand, Kirkley, and McConnell; Squires 1987a, b, c). Strand, Kirkley, and McConnell estimate a translog direct production function for the limited entry Atlantic surf clam fishery. They plot isoquants between input pairs to show that inputs are not used in fixed proportions. Squires estimates a translog profit function with three variable inputs (labor, fuel, and capital) for the open access New England otter trawl fishery. The elasticities of substitution he computes indicate considerable substitutability between input pairs. However, the fishing vessels in his work do not face input restrictions.

This paper argues that conventional elasticities of substitution, such as those used by Squires (1987c), cannot be used to evaluate a regulatory program that restricts the use by fishermen of certain inputs. These elasticities are obtained from a model that assumes that all inputs can be chosen freely by the fisherman. As an alternative, this paper proposes the calculation of the elasticity of intensity (Diewert), which describes the relationship between an unrestricted and a re-

Diane P. Dupont is an assistant professor, Department of Economics, Brock University, St. Catharines, Ontario. This paper was written while the author was an assistant professor, Department of Agricultural Economics and Business, University of Guelph.

The author is grateful for suggestions, comments, and support given by Gordon Munro, Phil Neher, and Bill Schworm. Additional help came from Erwin Diewert and Terry Wales. This paper has benefitted greatly from comments made by two anonymous reviewers, as well as Dale Squires, Jim Kirkley, Stephen Swallow, and, especially, Steven Renzetti. David Reid, and Paul McGillivray, in the Program, Planning, and Economics Branch, Department of Fisheries and Oceans, Pacific Region, Vancouver, British Columbia, Canada, have been generous in providing access to survey data and background material. All errors and opinions are the author's.

¹ Vessel tonnage is a commonly used measure of vessel size and is frequently chosen as the restricted input.

stricted input, to test for the degree of substitutability in a regulated industry subject to input controls. This elasticity may be defined within the context of the partial static equilibrium model discussed by Brown and Christensen. Under this model, fishing firms maximize restricted profits by choosing output (catch) and variable input quantities subject to the constraint that some inputs are restricted. Using this framework the paper shows that conventional elasticities of substitution are biased relative to elasticities of intensity.

The normalized quadratic restricted profit function (Diewert and Ostensoe) is used to estimate the fish harvest technology and the substitution possibilities for the British Columbia commercial salmon fishery, an important regulated Canadian fishery. The potential usefulness of the elasticity of intensity methodology to fisheries regulators is illustrated with the empirical results.

The next section considers the adoption of input restrictions in a fishery. Section three explains why using conventional elasticities of substitution to evaluate input control programs is inappropriate for a regulated fishery, or any regulated industry. Section four discusses the case study, estimation, and data, while section five reports results. The last section provides conclusions and suggestions to policy makers.

Input Restrictions in a Fishery

The conventional argument in favor of input restrictions is straightforward. If the fishery is open access, fishermen continue to enter the industry until total harvesting costs equal total revenues. Alternatively, if there are limited entry restrictions but no property rights to the fish, existing fishermen upgrade their vessels until the same equality is met. In both instances, fishery rent can be entirely dissipated (Gordon, Clark). The fish input is treated as having a zero price. Instead of charging fishermen for the use of this input, government regulators have chosen to restrict the fisherman's use of purchased inputs. If the fisherman has only one input at his disposal, then restricting the input's use will preserve fishery rent completely.

If inputs are substitutable, it is more difficult to restrict input usage effectively. Given an input restriction on vessel tonnage, a fisherman may reoptimize over the set of inputs under his control. For example, he may visit more and further-away fishing grounds, thereby increasing fuel consumption. Alternatively, he may fish

longer hours and use more labor as well as purchase more gear/equipment. In this way, he tries to increase output and increases his use of substitute inputs for the restricted input. Such direct substitution leads to higher harvest costs. Moreover, increased use of these substitutes leads in turn to increased use of their complements. Therefore, the use of variable inputs that may be direct complements to the restricted input may also increase. These actions, too, lead to higher harvesting costs and resource rent dissipation.

Substitution Possibilities between Unrestricted and Restricted Inputs

A partial static equilibrium model (Brown and Christensen) can be used to describe short-run behavior for firms subject to input restrictions. Assuming that the firm is free to vary the quantity of output, then a restricted profit function can be specified. The firm maximizes restricted profit by choosing the quantity of output supplied, y , and the quantities of variable or unrestricted inputs, $X = (x_1, x_2, \dots, x_n)$. In addition to facing exogenous input and output prices, the firm faces upper bounds on the use of inputs, $Z = (z_1, z_2, \dots, z_m)$.

Equation (1) gives the restricted profit function:

$$(1) \quad \pi^r(p_y, W, Z) = \max_{x,y} \{p_y \cdot y - W^T X; (y, X; Z) \in T\}.$$

The output price is denoted by p_y . The vector, W , represents market prices for the n variable inputs. When the restricted profit function in equation (1) fulfills a set of well-known properties defined in Diewert, it is dual to the underlying production function $F(X; Z)$ and to the production possibilities set, T .

Elasticity measures are obtained from the output supply and input demand equations which, in turn, are derived by differentiating the restricted profit function with respect to prices. Own- and cross-price elasticities of demand indicate the nature of the relationships between variable inputs. To determine whether variable inputs are complements or substitutes, it is sufficient only to know the cross-price elasticities of input demand. Equation (2) defines these elasticities, where x_i is the input demand function for the i th input. The elasticity indicates how the optimal employment of x_i by the firm will change in response to an incremental change in the price of the k th input.

$$(2) \quad \epsilon_{ik} = \frac{\partial x_i(p_y, W; Z)}{\partial W_k} \cdot \frac{w_k}{x_i}, \quad i, k = 1, 2, \dots, n.$$

Inputs are substitutes when ϵ_{ik} is positive and complements when it is negative.

A measure of substitution possibilities between unrestricted and restricted inputs is obtained with elasticities of intensity (Diewert).² Because these elasticities are derived from a restricted profit function, all prices and quantities of the restricted inputs, other than the one being considered, are held constant. Equation (3) defines these elasticities.

$$(3) \quad \xi_{ij} = \frac{\partial x_i(p_y, W; Z)}{\partial z_j} \cdot \frac{z_j}{x_i}, \quad i = 1, 2, \dots, n, \text{ and } j = 1, 2, \dots, m.$$

The Z_j 's are the actual levels of the restricted inputs. A negative elasticity shows a substitute relationship and a positive elasticity, a complementary one. For example, suppose that the value of the elasticity between tonnage and gear is -1.5 . This says that an incremental increase in tonnage, the restricted input, would result in a substantial reduction in the optimal employment of gear, an unrestricted input, by the fishing firm. Hence, the firm regards these two inputs as substitutes.

The relationship between the conventional cross-price elasticity of input demand and the elasticity of intensity can be established in the following way. To obtain the former elasticity, define a full static equilibrium profit function as one in which restricted inputs are treated as variable. Partition the input set into two parts, X and Z , and the input price set into two parts, W and R . The profit function becomes $\pi(p_y, W, R)$. Then, the cross-price elasticity of demand between inputs x_i and z_j (where z_j would be a restricted input in the restricted profit function) is given in equation (4):

$$(4) \quad \frac{\partial^2 \pi(p_y, W, R)}{\partial w_i \partial r_j} \cdot \frac{r_j}{x_i} = \frac{\partial x_i}{\partial r_j} \cdot \frac{r_j}{x_i}.$$

However, the right-hand side of (4) can be rewritten as in equation (5):

$$(5) \quad \frac{\partial x_i}{\partial r_j} \cdot \frac{r_j}{x_i} = \frac{\partial x_i}{\partial z_j} \cdot \frac{\partial z_j}{\partial r_j} \cdot \frac{r_j}{x_i}.$$

Equation (6) is obtained by multiplying both the numerator and denominator of the right-hand side

of equation (5) by z_j/z_j and rearranging the resulting expression,

$$(6) \quad \frac{\partial x_i}{\partial r_j} \cdot \frac{r_j}{x_i} = \left(\frac{\partial x_i}{\partial z_j} \cdot \frac{z_j}{x_i} \right) \left(\frac{\partial z_j}{\partial r_j} \cdot \frac{r_j}{z_j} \right), \text{ or}$$

$$\epsilon_{ij} = \xi_{ij} \cdot \epsilon_{jj}.$$

Equation (6) shows that the conventional cross-price elasticity of input demand between inputs i and j (obtained by assuming full static equilibrium and defining the input z_j as a variable input) is the product of the elasticity of intensity between inputs x_i and z_j and the own-price elasticity of z_j . If the own-price elasticity of demand for z_j is elastic, then the cross-price elasticity of demand will be greater than the corresponding elasticity of intensity.³ The converse is true if the own-price elasticity for z_j is inelastic.

Fishery Background, Characteristics, and Data

The commercial salmon fishery in British Columbia is the province's most lucrative and important fishery with annual revenues of approximately \$212 million in 1987 (and \$160 million in 1982, the year for which data are available to estimate the harvest technology). Salmon account for 40% of quantity and 70% of landed value of all catches (British Columbia Ministry of the Environment 1983, 1988). There are five species of salmon: sockeye, pink, chum, chinook, and coho. Four types of vessels are used, and there is a large degree of uniformity of species composition in the catches of the different vessel types. The vessel types are seine, gillnet, troll, and a combination of the last two called gillnet-troll. Each vessel type exhibits some unique characteristics, but all use some combination of labor, fuel, gear, and capital (vessel) inputs. Seine and gillnet vessels use a variety of nets to entrap the salmon, while troll vessels make use of lines, hooks, and bait to entice the salmon. Gillnet-troll vessels use both nets and troll lines. Seine vessels have the largest tonnage and employ from four to six people. Troll vessels are slightly smaller. Gillnet-troll and gillnet vessels are at the lower end of the spectrum. Each vessel is small enough relative to the total number of vessels—in 1982 this was 4,528—to preclude any one vessel exercising market power.

²Diewert defines the elasticity of intensity between a variable commodity (output or input) and a primary (fixed) input and uses the concept to provide comparative statics results for international trade theorems. The elasticity is reinterpreted here to explain the relationships between unrestricted and restricted inputs.

³This relationship explains why Squires (1987a) finds large elasticity values between capital and labor, e.g., 0.726, and capital and fuel, e.g., 2.125, when he estimates a full static equilibrium profit function. Capital's own-price elasticity is -2.821 . However, because conditions in the fishery that he studies meet the specifications of the full static equilibrium approach, the use of conventional elasticity measures is appropriate.

Since 1969 the salmon fishery has been subject to a limited entry program which sets an upper bound on the number of participating vessels. Since 1971, the regulator (Department of Fisheries and Oceans) has restricted the level of tonnage per vessel as a way of controlling fishing effort per vessel. To make the program effective, the tonnage of a replacement vessel can be no larger than that of the old vessel. Furthermore, pyramiding of vessels (i.e., combining the tonnage of two retiring vessels to apply to a new vessel) has been outlawed since 1980. These restrictions impose an effective upper bound upon the tonnage of a given vessel.

In order to examine the ability of fishermen in the British Columbia commercial salmon fishery to subvert regulations through increased use of unrestricted inputs in place of restricted inputs, this research postulates a restricted profit function for each of the four vessel types. Each vessel owner is assumed to maximize restricted profit over the season by choosing the quantity of catch and the quantities of three variable inputs (fuel, labor, and gear/equipment which includes nets, hooks, lines, etc.). The gear/equipment input is a key unrestricted input to which the firm may turn in response to restrictions upon tonnage (Pearse, Rettig). The constraints upon the fisherman's behavior are the harvest production function, the output price, and prices of variable inputs, and the existing quantities of three restricted inputs. These inputs are vessel tonnage and number of fishing days, both restricted by the regulating agency, and the stock of fish available for the season, restricted by nature.⁴

The research brings together a number of diverse data sets. The first, a cross-sectional random survey of Pacific Coast fishermen for 1982, gives fuel expenditures, vessel characteristics, labor quantities, number of fishing days, and gear/equipment inventory information by vessel.⁵ The second set of data is catch and revenue information collected for the regulator by processing companies. Prices for the five subspecies of salmon are used to calculate a Divisia index for aggregate output price and the asso-

ciated aggregate quantity index for each vessel.⁶ Opportunity cost wages are constructed for labor (Squires 1987b) by using average weekly earnings in an industrial composite category by region, based upon each vessel's homeport region.⁷ Marine fuel prices for eleven centers come from Esso Canada Limited. The gear/equipment input, consisting of nets, lines, etc., is considered to be a malleable capital good whose services are not exhausted in one year. A rental cost of gear is calculated for each component using Schworm's modification of Jorgenson's capital services price formula. Schworm's formula allows for the inclusion of the cost of repairs and maintenance, an important consideration for the fishing firm. Quantity and unit rental price data are used to construct a Divisia gear price index for each vessel.

License records from the Department of Fisheries and Oceans Canada give values for the restricted amount of tonnage per vessel. Data on catch and escapement in each of twenty-nine management areas are used to calculate an index of stock abundance. Because stock abundance varies across fishing grounds, each vessel can encounter a different quantity of fish. Stock encountered is calculated as the relative price-weighted abundance of each of the five salmon species in each fishing area weighted by the number of weeks of fishing in that area.

Restricted Profit Function Specifications and Estimation

The normalized, quadratic restricted profit function (Diewert and Ostensoe), one of a class of flexible functional forms that allows for non-constant elasticities of substitution, represents the objective function for the fishing firm. This function has two features that give it an edge

⁴ The number of days fished is a constrained value since the government uses fishing ground openings and closings for conservation purposes. Each vessel owner would likely fish more days if left on his own. Historical data show that, prior to strict regulations on fishing days, the average number of fishing days was much larger than currently (Pearse).

⁵ Only those vessels that caught salmon or salmon and less than 5% of incidental species such as sturgeon, shrimp, and ling cod have been chosen from the survey. The salmon-fishing fleet comprises the majority of vessels operating in the west coast fisheries. Vessels fishing the two other major species, roe herring and halibut, have not been included because they are regulated in ways that differ from regulation of the salmon fishery.

⁶ Insufficient data prevent the adoption of a multioutput specification defined over the five salmon species unless the inputs are aggregated to form a single index. Given that the issue in the paper is whether fishermen can substitute unrestricted for restricted inputs, it is considered more important to disaggregate inputs than outputs. (This is in contrast to Kirkley and Strand, who concentrate upon the relationships between outputs and aggregate their inputs into a single input index.) Aggregating over outputs allows one to examine the relationships between inputs or whether the harvest technology exhibits jointness in inputs (Squires 1987a). If more data were available, it would be interesting to estimate a model with both disaggregated outputs and inputs.

⁷ A share system whereby crew members receive a share of the value of the catch is used in the fishery, but the system is neither complete nor uniform. The system is used mostly for the seine fleet, and shares vary by species and vessel. The available data are incomplete with respect to the relevant values. As an alternative, the opportunity cost approach (Squires 1987b) is a reasonable substitute. A comparison of the data available on shares received by crew members with the calculated opportunity cost wage reveals that the latter follows very closely the former.

over the translog (Squires 1987a, b, c) and the generalized Leontief (Kirkley and Strand).⁸

First, to estimate the elasticity of intensity, a researcher must have parameter estimates for the cross-restricted input terms appearing in a restricted profit function. If the researcher were to adopt the translog, she would have two options. The first would be to estimate restricted input share equations, along with variable input share equations (McKay, Lawrence, and Vlastuin). However, this option requires that markets for restricted inputs are in equilibrium at current market rental prices. This situation is unlikely in a regulated industry. The second option would be to estimate the restricted profit function along with the variable quantity share equations. However, this option may introduce multicollinearity and degrees-of-freedom problems because all parameters appear in the restricted profit function, whereas each input share equation has less than the complete set of parameters. Depending upon the number of variable inputs/outputs and restricted inputs specified in a model, a restricted profit function can often have more than thirty parameters that need to be estimated.

Alternatively, using the normalized quadratic functional form allows the researcher to retrieve the required parameter estimates by estimating a system of output supply/unrestricted input demand equations. There is no need to assume equilibrium in markets for restricted inputs. Furthermore, because each equation has less than the complete set of parameters, there is less likelihood of data-based multicollinearity problems or matrix-singularity problems leading to nonconvergence of parameter estimates.

The second advantage of the normalized quadratic is that a researcher can impose convexity in prices on parameter estimates of the restricted profit function and continue to identify separate elasticities between individual pairs of inputs. If curvature conditions are satisfied, estimated demand and supply functions will be well-behaved. This property is desirable for simulating rents obtained under alternative regulatory schemes (Dupont). Previous estimates of fish harvesting technology have not met assumed

curvature conditions (Squires 1987a, b, c; Kirkley and Strand).⁹

The normalized quadratic restricted profit function defined over four variable quantities (one output and three variable inputs) and three restricted inputs is given in equation (7).¹⁰

(7)

$$\begin{aligned} \pi'(P, W, Z) = & 1/2 \left(\sum_j \alpha_j z_j \sum_i \sum_k a_{ik} p_i p_k \right) / p_y \\ & + 1/2 \left(\sum_i \beta_i p_i \sum_j \sum_h b_{jh} z_j z_h \right) / z_s \\ & + \sum_i \sum_j c_{ij} p_i z_j + \left(\sum_i \beta_i p_i \sum_j b_{ij} z_j \right) / z_s \\ & + 1/2 \left(\sum_i b_{0i} \beta_i p_i \right) / z_s + \sum_i c_i p_i. \end{aligned}$$

Variable quantity prices are indexed by i, k in the following order: p_y (price of output, i.e., salmon catch) p_l (price of labor), p_f (price of fuel), and p_g (price of gear and equipment). Indexing of fixed quantities, j, h is as follows: z_s (stock of fish), z_t (tonnage), z_d (number of fishing days). Because the function is normalized, *numéraires* of p_y and z_s are chosen. Define the matrix A with elements a_{ik} . Because of the linear relationships between rows and columns in the matrix A caused by linear homogeneity, the first row and column of A , e.g., a_{yk} through a_{yg} ($k = y, l, f, g$) are taken to be vectors of zeroes.

The parameters to estimate are $\alpha_j, a_{ik}, \beta_i, b_{jl}, c_{ij}, b_j, c_i, b_0$. Diewert and Wales note that the α_j ($j = s, t, d$) may be chosen arbitrarily and suggest they be set equal to $1/z'_j$, where z'_j is the fixed factor vector for the first observation. Likewise, the β_i ($i = y, l, f, g$) may be set equal to $1/p'_i$. This convention is adopted here.

The normalized quadratic restricted profit function described in equation (7) satisfies the conditions required for it to represent the underlying production technology. The function is linearly homogenous in prices. Symmetry in cross-price terms is obtained by defining the matrix A to be symmetric. The restricted profit function satisfies convexity in variable quantity

⁸ Because the normalized quadratic is a relatively new flexible functional form, there has not yet been any empirical or theoretical work (like that of Christensen and Caves and Barnett and Lee, who find that translog elasticities tend to be closer to one in absolute value, whereas Leontief elasticities tend to gravitate toward zero) to compare it to the translog or the generalized Leontief. Diewert and Wales, however, compare elasticity values obtained from cost functions represented by a translog, a generalized Leontief, and a new functional form that is similar to the normalized quadratic (the symmetric generalized McFadden). Their findings agree with those of the authors cited above. In addition, they find that elasticities from the symmetric generalized McFadden tend to fall between the values for the two other functional forms.

⁹ Wales suggests that nonconvexity may indicate that the chosen functional form does not provide a good fit for the data over the sample range used. Some reasons for failing to obtain convexity are insufficient price variation in the data, multicollinearity, and aggregation of input or output quantities to obtain indexes.

¹⁰ Lopez shows that the function imposes homothetic output-input separability upon the harvest technology. This means that the marginal rates of substitution between input pairs are independent of the levels of individual outputs. Because a single output framework is necessitated by the data, this restriction is a minor limitation of this function.

prices globally (and locally) whenever the A matrix is positive semidefinite.

Instead of estimating the restricted profit function in (7), it is more convenient to estimate the system of four variable quantity equations given in (8)–(9). These equations, one for each of output and the three variable inputs (fuel, labor, and gear/equipment), are obtained from (7) by using Hotelling's lemma. These equations are formulated in actual quantities, not input or revenue shares; therefore, all four equations must be estimated to obtain the parameters of interest.

$$\begin{aligned}
 (8) \quad \frac{\partial \pi^r}{\partial p_y} &= y^*(p_y, p_l, p_f, p_g; z_s, z_l, z_d) \\
 &= -1/2 \left(\sum_j \alpha_{yz_j} \sum_i \sum_k a_{ik} p_i p_k \right) / p_y^2 \\
 &\quad + 1/2 \left(\beta_y \sum_j \sum_h b_{jh} z_j z_h \right) / z_s + \sum_j c_{yz_j} \\
 &\quad + \beta_y \left(\sum_j b_{z_j} \right) / z_s + (1/2 b_0 \beta_y) / z_s + c_y; \\
 (9) \quad \frac{\partial \pi^r}{\partial p_{xi}} &= -x_i^*(p_y, p_l, p_f, p_g; z_s, z_l, z_d) \\
 &= \left(\sum_j \alpha_{yz_j} \sum_k a_{ik} p_k \right) / p_y \\
 &\quad + 1/2 \left(\beta_i \sum_j \sum_h b_{jh} z_j z_h \right) / z_s \\
 &\quad + \sum_j c_{iz_j} + \beta_i \left(\sum_j b_{z_j} \right) / z_s \\
 &\quad + (1/2 b_0 \beta_i) / z_s + c_i \quad \text{all } i = l, f, g.
 \end{aligned}$$

Cross-equation and symmetry restrictions have already been imposed in equations (8)–(9).¹¹ These restrictions include $a_{ik} = a_{ki}$ in all equations for ik and $b_{jh} = b_{hj}$ in all equations and for each jh .

If convexity is rejected by the data, it can be imposed by a reparameterization of the A matrix (Wiley, Schmidt, and Bramble). This reparameterization uses the product of a matrix D and its transpose to replace the A matrix, i.e., $A = DD^T$. The D matrix is a lower triangular matrix with zeroes in the first column as shown in equation (10). (Recall that the first row and column of the A matrix are zeroes due to the linear relationships between the rows and columns imposed by linear homogeneity.)

$$(10) \quad \begin{bmatrix} a_{yy} & a_{yl} & a_{yf} & a_{yg} \\ a_{yl} & a_{ll} & a_{lf} & a_{lg} \\ a_{yf} & a_{lf} & a_{ff} & a_{fg} \\ a_{yg} & a_{lg} & a_{fg} & a_{gg} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ d_1 & 0 & 0 & 0 \\ d_2 & d_2 & 0 & 0 \\ d_4 & d_5 & d_6 & 0 \end{bmatrix} \begin{bmatrix} 0 & d_1 & d_2 & d_4 \\ 0 & 0 & d_3 & d_5 \\ 0 & 0 & 0 & d_6 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

While it is still possible to obtain separate elasticity estimates for each pair of inputs, the reparameterization requires a nonlinear estimation technique. A new set of equations must be estimated using a nonlinear maximum likelihood procedure because the a_{ik} parameters are replaced by the appropriate combinations of d parameters from the D matrix. The correspondence between the a_{ik} and d parameters is as follows: $a_{ll} = d_1^2$, $a_{lf} = d_1 d_2$, $a_{lg} = d_1 d_4$, $a_{ff} = d_2^2 + d_3^2$, $a_{fg} = d_2 d_5 + d_2 d_4$, and $a_{gg} = d_4^2 + d_5^2 + d_6^2$.

Prior to estimation, additive disturbance terms are appended to each of the four variable quantity equations in (8)–(9). Estimation begins with the linear system of equations (8)–(9). Zellner's iterative technique for seemingly unrelated regressions is used in four separate systems of equations (one for each of four vessel types). [The number of observations is as follows: seine (21), gillnet (80), troll (84), and gillnet-troll (60).] Resulting parameters are checked for acceptance of convexity in prices. Because estimates for one of the vessel types (troll) are consistent with this characteristic, no further estimation of this sample is required. The other three samples (seine, gillnet, and gillnet-troll) do not accept convexity.¹² Therefore, in order to impose convexity, these samples are reestimated using maximum likelihood with a system of four nonlinear equations. Space limitations preclude reporting parameter and elasticity estimates for all four samples. (The author will supply them upon request.) Tables 1 and 2 report parameter estimates, their standard errors, and other summary statistics for two representative samples, the seine and the gillnet-troll. Calculated elasticities of intensity are found to be invariant to the imposition of convexity in prices. This is not surprising since only the cross-price terms are replaced when convexity is imposed.

¹¹ At a 5% level of significance, the seine and gillnet-troll samples did not reject symmetry.

¹² This finding manifests itself in downward-sloping output supply functions. Squires (1987c) finds this is the case, as well. Output aggregation is one possible cause; however, insufficient price variation (often a problem with cross-sectional data) is a more likely one. In order to examine the impact of output aggregation, further estimation is undertaken with two outputs formed by aggregating over fresh and canned salmon species. However, convexity is still rejected by the same samples as previously. Squires (1987c) obtains the same result for an even finer disaggregation of outputs.

Table 1. Restricted Profit Function Parameter Estimates—Seine

Variable Name	Coefficient Value	Standard Error	Variable Name	Coefficient Value	Standard Error
d_1	-0.139 ^a	0.060	c_{yd}	-1.670 ^a	0.721
d_2	-0.272 ^b	0.151	c_y	1.220	1.953
d_3	-0.101E - 05	0.308	c_{ly}	0.199	0.661
d_4	0.323 ^a	0.106	c_{ly}	1.473	1.365
d_5	0.665E - 06	0.232	c_{ld}	-2.413 ^a	0.614
d_6	0.601E - 07	0.124	c_l	-2.478	1.988
b_{ll}	3.498	2.696	c_{fs}	0.189	1.153
b_{ld}	-3.728 ^b	2.283	c_{fd}	-1.414	2.023
b_{dd}	-1.424 ^a	0.704	c_{fd}	-2.999 ^a	0.949
b_0	-3.581	3.001	c_f	-1.030	2.625
b_l	-0.418	3.040	c_{gs}	13.711 ^a	7.744
b_d	5.411 ^a	1.103	c_{gd}	19.771 ^a	10.877
c_{ys}	-0.013	0.708	c_{gd}	-9.420 ^a	5.979
c_{yt}	1.678	1.538	c_g	-31.171 ^a	13.713

^a Five percent significance.^b Ten percent significance.**Table 2. Restricted Profit Function Parameter Estimates—Gillnet-Troll**

Variable Name	Coefficient Value	Standard Error	Variable Name	Coefficient Value	Standard Error
d_1	0.128	0.111	c_{yd}	0.453 ^a	0.236
d_2	-0.127	0.245	c_y	-0.370	1.436
d_3	-0.293 ^a	0.127	c_{ly}	-0.033	0.187
d_4	0.591E - 02	0.028	c_{ly}	-1.566 ^a	0.569
d_5	0.572E - 02	0.029	c_{ld}	-0.260 ^a	0.096
d_6	0.145E - 08	0.040	c_l	0.792	0.663
b_{ll}	-2.681 ^a	1.557	c_{fs}	-0.156	0.197
b_{ld}	0.458 ^a	0.204	c_{fd}	-1.479 ^a	0.594
b_{dd}	0.073 ^b	0.053	c_{fd}	-0.364 ^a	0.104
b_0	-3.490	1.910	c_f	1.161	0.682
b_l	3.038 ^a	1.618	c_{gs}	40.551 ^a	22.139
b_d	-0.567 ^a	0.216	c_{gd}	-80.144 ^b	61.727
c_{ys}	0.273	0.558	c_{gd}	3.279 ^a	1.097
c_{yt}	2.697 ^a	1.267	c_g	-18.401	68.517

^a Five percent significance.^b Ten percent significance.

Results

Results for the seine and gillnet-troll fleets are presented because these vessel types represent two extremes in the fishery and, therefore, establish bounds on results for the entire fleet. For the two vessel types, tables 3 and 4 report estimated cross-price elasticities for variable input pairs. These elasticities are calculated using parameter estimates from the nonlinear estimation procedure and mean values from each sample for the regressors. The appendix gives formulas from the normalized quadratic restricted profit function for these elasticities and for the elasticities of intensity.

Elasticities of intensity are used to examine the relationships between unrestricted and restricted inputs. Table 5 presents these elasticities

Table 3. Cross-Price Elasticity Estimates for Seine Vessels

Quantities	Labor ^a	Fuel	Gear
Prices			
Labor	-0.025 ^b (0.021)	-0.050 ^c (0.029)	0.024 ^c (0.014)
Fuel	-0.055 ^c (0.032)	-0.111 (0.123)	0.055 (0.044)
Gear	0.013 ^c (0.007)	0.025 (0.020)	-0.012 ^d (0.008)

^a Elasticity estimates use means of the data.^b Standard errors are in parentheses. Asymptotic standard errors use the formula for the variance of a random variable that is a nonlinear function of several random variables (Kmenta).^c Five percent significance.^d Ten percent significant.

Table 4. Cross-Price Elasticity Estimates for Gillnet-Troll Vessels

Quantities	Labor ^a	Fuel	Gear
Prices			
Labor	-0.044 ^b (0.076)	0.060 (0.107)	-0.001 (0.007)
Fuel	0.078 (0.140)	-0.677 ^c (0.277)	0.009 (0.025)
Gear	-0.00006 (0.0003)	0.003 (0.0007)	-0.4E - 6 (0.2E - 5)

^a Elasticity estimates use means of the data.^b Standard errors are in parentheses. Asymptotic standard errors use the formula for the variance of a random variable that is a nonlinear function of several random variables (Kmenta).^c Five percent significance.^d Ten percent significance.**Table 5. Elasticity of Intensity Estimates**

Restricted Inputs	Seine Vessels ^a		Gillnet-Troll Vessels ^b	
	Tonnage	Days	Tonnage	Days
Variable quantities				
Labor	-0.010 (0.029)	0.004 (0.051)	0.452 ^c (0.170)	0.277 ^c (0.065)
Fuel	0.366 ^c (0.180)	0.167 (0.237)	0.571 ^d (0.356)	0.507 ^c (0.129)
Gear	-0.425 ^d (0.251)	0.406 (0.336)	1.608 (1.253)	-0.115 (0.389)

^a Elasticity estimates use means of the data.^b Standard errors are in parentheses. Asymptotic standard errors use the formula for the variance of a random variable that is a nonlinear function of several random variables (Kmenta).^c Five percent significance.^d Ten percent significance.

ties for the seine and gillnet-troll vessels for both inputs restricted by the regulator (tonnage and fishing days). Looking first at seine vessel elasticities associated with tonnage, labor and gear are substitutes for the restricted input, while fuel is a complement. The elasticity values are large (especially relative to the cross-price elasticities) and significant for the fuel and gear inputs. The direct substitution effect discussed above means that rent dissipation occurs as seine vessel-owners substitute toward the use of increased amounts of gear and labor inputs. The indirect effect of the increased use of labor means an increased use of fuel, its complement. Because gear and fuel are substitutes, the net impact depends upon the relative strength of the sets of elasticities. Nonetheless, the results suggest rent dissipation by seine vessels. This finding accords with Pearse's observations and Dupont's estimates; it is of concern because seine vessels, although smallest in absolute numbers

(8% of the fleet), dominate the fishery by percentage of fish landed (38% in 1982).

Results for the gillnet-troll fleet suggest that all three variable inputs are complements for tonnage (table 5). Furthermore, the elasticity values are large, ranging from 0.452 (labor) to 1.608 (gear), and significantly different from zero in two out of three cases. Thus, there is no direct substitution effect. Moreover, although the cross-price elasticities show that labor and gear are complementary inputs, the values are extremely small and not significantly different from zero. The results support the hypothesis that the gillnet-troll vessel does not have many ways of dissipating fishery resource rent in the face of an input restriction upon a vessel's use of tonnage.

The conclusions are reversed when the elasticities of intensity associated with restricted fishing days are examined (table 5). (Recall that each elasticity is a partial elasticity obtained by holding constant the quantities of other restricted inputs.) With one exception, each variable input has a complementary relationship with fishing days in the samples. The values are inelastic but generally significantly different from zero. Hence, the fishing days restriction is effective in preventing input substitution that can lead to rent dissipation. The one exception is the ability of vessel-owners in the gillnet-troll sample to substitute more gear against a fishing days restriction. They do this by purchasing both gillnet and troll gear/equipment. Then, when a fishing ground is closed to one type of gear, the gillnet-troll vessel-owner switches to the alternative unrestricted type of gear.

Elasticities, not reported in this paper, for the other two vessel types show that the troll vessel is strikingly similar to the seine vessel in its ability to exploit substitution possibilities, while gillnet-troll results represent the gillnet vessel well (with respect to tonnage elasticities). Thus, the British Columbia limited entry program with tonnage restrictions per vessel appears to have been moderately successful in preventing only two of the four vessel types from dissipating resource rent in the salmon industry. The inability of fishermen to find substitutes for the restricted fishing days input has undoubtedly contributed to the regulator's limited success.

Conclusions

Fisheries regulators have now had two decades of experience with quantitative input restrictions

imposed to preserve resource rent. This paper examines whether input control programs can be successful by asking whether the fishing harvest technology allows fishermen to subvert the intentions of the programs by substituting unrestricted for restricted inputs. The paper confirms for the British Columbia commercial salmon fishery, an important regulated Canadian fishery, that these programs have not been effective, as many have argued (Pearse, Rettig). Examination of the elasticities of intensity between the restricted input, vessel tonnage, and the unrestricted inputs provides evidence about the ability of two of the four vessel types (seine and troll) in the fishery to substitute unrestricted for restricted inputs. In so doing, the owners of seine and troll vessels are able to dissipate fishery rent (Dupont). Restrictions on fishing days, on the other hand, do not permit much substitution and are better choices for regulated inputs.

In light of these findings, it is appropriate to argue that the regulators should consider phasing out tonnage control schemes, increasing the use of fishing days restrictions, or replacing input controls entirely with individual transferable vessel quotas or royalty taxes. Difficulties with the imposition of a royalty tax include the determination of the correct rate, the necessity of an annually fluctuating rate in response to stock variability, disincentive effects that encourage illegal landings, and possible distributional effects of tax incidence. Furthermore, taxes are not popular with fishermen. On the other hand, the individual transferable vessel quota (ITVQ) is generating a great deal of interest among fisheries economists (Christy, Scott). One problem with adopting an ITVQ is that it is tantamount to the creation of property rights. Furthermore, as New Zealand's recent experience shows, the ITVQ can create additional problems for regulators. These problems include the bycatch problem (if the fisherman does not have a quota for a particular species, he is likely to throw any unwanted species caught back overboard where there is only a slim chance of survival), administration of quotas associated with a fluctuating fish stock, and what to do with fishermen who catch more than their quotas (Clark, Major, and Mollett).

Given the problems with the alternative forms of regulation and also given inertia in the fishery, input controls likely will continue to be used. However, the potential resource rent available in fisheries is so great that regulators cannot afford to incur the opportunity cost imposed by inefficient regulation. This paper's approach and

findings will be useful to regulators of other fisheries or industries charged with finding innovative ways of implementing input control programs. The regulator would need only to examine the elasticities of intensity between various pairs of variable and restricted inputs. Those inputs with few substitutes or with positive or zero elasticities of intensity would be good choices for input restriction programs. Alternatively, if the regulator chooses to adopt a set of input restrictions, an examination of pairs of elasticities of intensity would help to determine the best mix of restricted inputs. This represents an improvement over the haphazard or politically motivated choices usually made for controlled inputs.

[Received November 1989; final revision received April 1990.]

References

- Barnett, W. A., and Y. W. Lee. "The Global Properties of the Mini-flex Laurent, Generalized Leontief, and Translog Functional Forms." *Econometrica* 53(1985): 1421-37.
- Brown, R. S., and L. R. Christensen. "Estimating Elasticities of Substitution in a Model of Partial Static Equilibrium: An Application to U.S. Agriculture, 1947 to 1974." *Modelling and Measuring Natural Resource Substitution*, ed. E. Berndt and D. Field. Cambridge MA: MIT Press, 1979.
- British Columbia, Ministry of the Environment. Marine Resources Branch. *Fisheries Production Statistics of British Columbia 1982*. Victoria, 1983.
- . *Fisheries Production Statistics of British Columbia 1987*. Victoria, 1988.
- Christensen, L. R., and D. W. Caves. "Global Properties of Flexible Functional Forms." *Amer. Econ. Rev.* 70(1980):322-32.
- Christy, F. T., Jr. "Alternative Arrangements for Marine Fisheries: An Overview." Washington DC: Resources for the Future, Pap. No. 1, 1973.
- Clark, C. W. *Bioeconomic Modeling and Fishery Management*. New York: John Wiley & Sons, 1985.
- Clark, I. N., P. J. Major, and N. Mollett. "Development and Implementation of New Zealand's ITQ Management System." *Marine Res. Econ.* 5(1988):325-49.
- Diewert, W. E. "Applications of Duality Theory." *Frontiers in Quantitative Analysis*, ed. M. D. Intriligator and D. A. Kendrick. Amsterdam: North-Holland Publishing Co., 1974.
- Diewert, W. E., and L. Ostensoe. "Flexible Function Forms for Profit Functions and Global Curvature Conditions." University of British Columbia Disc. Pap. No. 87-06, 1987.
- Diewert, W. E., and T. J. Wales. "Flexible Functional Forms and Global Curvature Conditions." *Econometrica* 55(1987):43-68.

- Dupont, D. P. "Rent Dissipation in Restricted Access Fisheries." *J. Environ. Econ. and Manage.*, in press.
- Jorgenson, D. W. "Capital Theory and Investment Behaviour." *Amer. Econ. Rev.* 53(1963):247-359.
- Karpoff, J. M. "Suboptimal Controls in Common Resource Management: The Case of the Fishery." *J. Polit. Econ.* 95(1987):179-94.
- Kirkley, J. E., and I. E. Strand. "The Technology and Management of Multi-species Fisheries." *Appl. Econ.* 20(1988):1279-92.
- Kmenta, J. *Elements of Econometrics*. New York: Macmillan Co., 1977.
- Lopez, R. "Structural Implications for a Class of Flexible Functional Forms for Profit Functions." *Int. Econ. Rev.* 26(1985):593-601.
- McKay, L., D. Lawrence, and C. Vlastuin. "Profit, Output Supply and Input Demand Functions for Multiproduct Firms: The Case of Australian Agriculture." *Int. Econ. Rev.* 24(1983):323-39.
- Pearse, P. H. *Turning the Tide: A New Policy for Canada's Pacific Fisheries*. Commission on Pacific Fisheries Policy, final report. Ottawa: Supply and Services, 1982.
- Rettig, R. B. "License Limitation in the United States and Canada: An Assessment." *N. Amer. J. Fisheries Manage.* 4(1984):231-48.
- Schworm, W. E. "User Cost and the Demand for Capital." Discussion Paper 77-22. The University of British Columbia Disc. Pap. No. 77-22, 1977.
- Scott, A. D. "Development of Economic Theory on Fisheries Regulation." *J. Fisheries Rev. Board of Can.* 36(1979):725-41.
- Squires, D. "Fishing Effort: Its Testing, Specification and Internal Structure in Fisheries Economics and Management." *J. Environ. Econ. and Manag.* 14(1987a):268-82.
- . "Long-Run Profit Functions for Multiproduct Firms." *Amer. J. Agr. Econ.* 69(1987b):558-69.
- . "Public Regulation and the Structure of Production in Multiproduct Industries: An Application to the New England Otter Trawl Industry." *Rand J. Econ.* 18(1987c):232-48.
- Strand, I. E., J. Kirkley, and K. McConnell. "Economic Analysis and the Management of Atlantic Surf Clams." *Economic Analysis for Fisheries Management Plans*, ed. L. G. Anderson. Ann Arbor MI: Ann Arbor Science Publishing, 1981.
- Wales, T. "On the Flexibility of Flexible Functional Forms: An Empirical Approach." *J. Econometrics* 5(1977):183-93.
- Warming, J. "Om Grundrente af Fiskegrunde." (On Rent of Fishing Grounds," tr. P. Andersen (1983) in *History of Political Economy*). *Nationalokonomisk Tidsskrift* 49(1911):499-505.
- Wiley, D. E., W. H. Schmidt, and W. J. Bramble. "Studies of a Class of Covariance Structures." *J. Amer. Statist. Assoc.* 68(1973):317-23.
- Zellner, A. "An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias." *J. Amer. Statist. Assoc.* 57(1962):348-68.

Appendix

The own- and cross-price elasticities between variable inputs are given below in equation (A1) for the normalized quadratic restricted profit function.

$$(A1) \quad \epsilon_{ii} = \left(a_{ii} \sum_j \alpha_j \frac{z_j}{p_j} \right) \cdot \frac{p_i}{x_i} \quad \text{for } i = l, f, g.$$

The elasticity of intensity between tonnage and each of the variable inputs is given in equation (A2) for the normalized quadratic restricted profit function.

$$(A2) \quad \xi_{ij} = \left(\alpha_i \sum_k \alpha_k \frac{p_k}{p_y} + \frac{(B_i)}{z_i} (b_{iz_i} + b_{id}z_d + b_i) + c_{ii} \right) \cdot \frac{z_i}{x_i} \quad \text{for } i, k = l, f, g.$$

Efficient Spatial Allocation of Irrigation Water

Ujjayant Chakravorty and James Roumasset

In the presence of conveyance losses, the efficient quantity of water applied falls with distance from the water source, but the amount of water "sent" (including conveyance losses) actually increases with distance from the source, except toward the tail end of the irrigation system. This implies that if marginal cost pricing were implemented, farmers at the middle and lower reaches of the system would have to pay more money for less water received. The model is illustrated and alternative financing schemes compared for an empirically derived demand function for irrigation water.

Key words: benefit taxation, conveyance losses, irrigation, spatial efficiency, water.

Despite massive public investments in irrigation infrastructure, *ex post* evaluations of irrigation projects in developed and developing countries indicate that actual benefits are substantially below projected levels. Considerable evidence suggests that these low benefits are largely the result of poor on-farm water use efficiencies and rent-seeking activities that result from water charges that are low and often unrelated to water use (Chaudhry, Repetto, Bowen and Young). For example, farmers near the system headworks are said to consume a disproportionate share of irrigation water, while tail farmers are left with scanty and unreliable residual supplies (Reidinger, Wade).

Many governments, faced with increasing political pressure to conserve water and reduce fiscal deficits are considering higher water charges to decrease waste and increase cost recovery from project beneficiaries. There is also increased awareness that low water charges and loosely enforced water rationing guidelines lead to environmental damages and excessive mining of groundwater resources. In general, the problems of water allocation and low user charges contribute to derivative problems in achieving efficiency, equity, fiscal stability, and environmental sustainability.

In order to improve irrigation performance and promote sustainable use of agricultural water, an analytical framework for allocating water and levying water charges is needed (Repetto). Although water financing systems vary widely, traditional concepts of irrigation management imply that farmers be charged uniform prices for equal amounts of water delivered (Bishop and Long, Yoo and Busch, Burness and Quirk). Similarly, evaluations of irrigation systems implicitly assume that equal water allocation is desirable and that head versus tail disparities in water allocation are *prima facie* evidence of inefficiency and inequity. As shown below, however, irrigation policies that prescribe equal allocations of water to farmers may be inefficient because conveyance costs of water increase with distance from the source.

In this paper, conditions for efficient spatial allocation of irrigation water are specified that take into account conveyance losses caused by seepage, percolation and evaporation.¹ In the following section, a theoretical model is outlined that derives rules for optimal allocation of water supplied to farmers at various distances from a water source. Optimal water prices are developed at each location, and their implications for rents and equity are examined. The subsequent section illustrates the analytical results for an empirically derived water demand function.

The authors are, respectively, an acting assistant professor of agricultural and resource economics and a professor of economics, University of Hawaii.

This is Hawaii Agricultural Experiment Station Journal Series No. 3432.

The authors are grateful to Ted Bergstrom and Sumner La Croix for helpful discussions and to three anonymous referees of this *Journal* for comments that substantially improved the paper.

¹ Tolley and Hastings incorporate conveyance losses to investigate the gains from reallocation of water from the North Platte River. However, their model does not yield general results. The model proposed here is not only more general but shows the relationship between conveyance efficiency and irrigation system design, user charges, and farmer welfare.

Efficiency Rules: A Framework

Abstracting from the costs of information and enforcement, efficient spatial allocation requires equal marginal value products of water measured at a common source point (O'Mara). The purpose of this section is to derive the equal (gross) marginal product rule and demonstrate its implications for (net) marginal value products and water allocations at the farm level and for determination of optimal system size.

Consider a simplified one-period (one season) model of an irrigation system with water supplied from a point source to a canal. Seasonal variabilities in water supply and storage are not considered. Farmers draw water at various points along the canal located at a variable distance y from the source, $y = 0$ representing the water source, and y increasing away from it. Let $Q(y)$ be the gross volume of source water sent to a farmer located at y , and $q(y)$ be the net water received after conveyance losses.² Then the relationship between source water $Q(y)$ and received water $q(y)$ is given by

$$\begin{aligned} (1) \quad & q(y) = Q(y)h(y) \\ & q(y), Q(y) \geq 0; \\ & 0 \leq h(y) \leq 1; h(0) = 1; \\ & h'(y) < 0; h''(y) \leq 0, \end{aligned}$$

where $h(y)$ denotes conveyance efficiency as a function of distance of farmer from source. Conveyance efficiency decreases with distance at an increasing rate. The major sources of conveyance loss are seepage and percolation. Evaporation losses are considered unavoidable and are ignored because, even if the water is stored and conveyed, it will evaporate, although at a different rate. Investments in canal lining and maintenance reduce $h(y)$.

In this model, we abstract from differential conveyance investments over space and from the choice of conveyance technology.³ Conveyance losses from the farmgate to the root zone of the crop can be regarded as given such that the yield-water relationship is well defined in terms of water delivered at the farmgate. The model also abstracts from the externality effects of convey-

ance losses on downstream water quality or on irrigation return flows.

We assume a one-input, one-output production function for received water that holds for each farmer in the system.⁴ Output per unit area is then given by a production function $f(q)$, which has the usual properties that apply to stage II of the neoclassical production function: $f(\cdot) > 0$; $f'(\cdot) > 0$; $f''(\cdot) < 0$. The value of marginal product function for received water on land of a uniform quality can be written as

$$(2) \quad VMP_r = Pf'(q),$$

where P is the competitive crop price. From (2), $\partial VMP_r / \partial q < 0$, a result that follows from the concavity of the production function.

The value of marginal product for received water at the source (VMP_s) can be defined (by assuming the necessary continuity and differentiability properties and using the chain rule) as

$$(3) \quad VMP_s(y) = Pf'(Q) = Pf'(q)h'(y) = VMP_r h'(y)$$

from (1) and (2), and

$$(4) \quad \partial VMP_s / \partial y = Pf'(q)h''(y) < 0,$$

using (1). Condition (3) gives the relationship between the two functions VMP_r and VMP_s , one as a function of water measured at the farm and another as a function of water measured at the source. From (4), the value of marginal product of each unit of source water decreases with distance of the farm from source. Intuitively, as y increases, more source water must be sent to produce one unit of received water on the farm, leading to a lower marginal value product at the source.

VMP_s curves can now be derived as a function of source water for farmers at different locations along the canal. Define q^m and $Q^m(y)$ as the amounts of water at which $VMP_r = 0$ and $VMP_s(y) = 0$, respectively. Note that from (2), q^m is not a function of y . For the farmer at the source ($y = 0$), $q^m = Q^m(0)$, and when $y > 0$, $q^m = Q^m(y)h(y)$, from which we get $Q^m(0) < Q^m(y)$. In other words, the amount of source water at which VMP_s becomes zero increases with distance.

² The theoretical construct used in this paper parallels that presented in Caswell and Zilberman, although the nature of the problem is different.

³ For a model that allows for endogenous choice of conveyance technology and investment in on-farm water conservation, see Chakravorty, Hochman, and Zilberman.

⁴ The monocropping assumption is retained for model simplicity but can be relaxed easily by specifying a unique production function for each farmer. Differences in farmers' preferences, skills, and financial endowments can also be accommodated by indexing on the production function.

The above results can be used to draw $VMP_s(y)$ curves at different distances from the source as shown in figure 1. For the farmer located at the source, $VMP_s(0) = VMP_r$ and $q^m = Q^m$. Curves $VMP_s(y_1)$, and $VMP_s(y_2)$ at increasing distances from the source flatten out toward the x -axis, as shown.

The rule for optimal spatial allocation of water is now derived by maximizing consumers' plus producers' surplus, subject to a capacity constraint as follows:

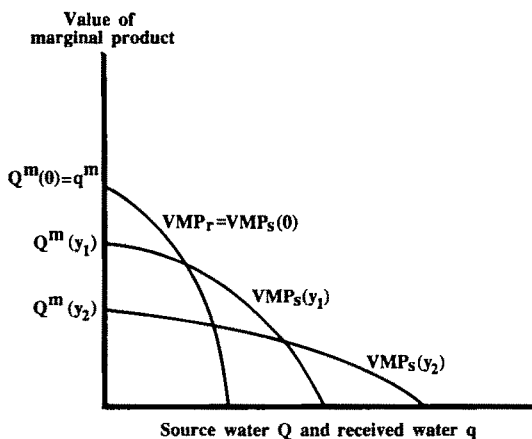
$$(5) \quad \text{Maximize} \quad \sum_{i=1}^n \int_0^{Q^i} VMP_s^i(\beta) d\beta - \int_0^Z C'(\theta) d\theta,$$

subject to

$$Z = \sum_{i=1}^n Q^i,$$

where i represents the i th farmer and $i = 1, 2, \dots, n$; Z is the total stock of water at source; $C'(Z)$ is the total long-run marginal cost of water; and β and θ are variables of integration. In this discrete optimization problem, each farmer i is associated with a distance y . In a continuous model, the same result could be obtained by summing the area under the VMP_s function for each y .

The first term in (5) represents the aggregate willingness-to-pay schedule. It is obtained by the horizontal summation of the VMP_s curves for each



Note: Q and q are related by the equation $q(y) = Q(y)h(y)$.

Figure 1. Position of value of marginal product curves at various distances from the water source where $0 < y_1 < y_2$

of the n farmers in the system. However, all the n farmers need not be within the efficient boundary of the system because the allocation rule derived from the above optimization exercise gives an efficient boundary beyond which farmers will not receive any water allocation. Choosing a value of n large enough will ensure the optimal solution (the allocation to some farmers may be zero).

$C'(Z)$ is the sum of the marginal costs of supply and distribution. The marginal cost of water supply includes the per-period equivalent of capital construction costs and costs of operation and maintenance of the head works. The marginal cost of distribution includes construction, operation, and maintenance of the canal as well as pumping and metering costs. The marginal cost of supply (distribution) usually decreases (increases) with capacity (Koenig). The resulting total marginal cost function could be rising or falling with Z . We have assumed that the topographical constraints on system expansion are not binding such that economies of scale in the total marginal cost schedule have been fully exploited, giving us a rising marginal cost curve. In an *ex ante* setting, the long-run marginal cost of water will be determined endogenously by choosing an optimal water stock, Z . However, if the model is applied to an existing project with a fixed capacity, $C'(Z)$ will be a short-run marginal cost function.

Homogeneity in demand is assumed for the purposes of allocation of capital costs among consumers. The marginal cost-based pricing rules, however, generally will not be optimal if users are heterogenous or if there are decreasing costs or jointness of supply. In those situations, other pricing rules, such as those based on incremental pricing, might be more relevant (Loehman and Whinston).

From (5), the following lagrangian can be maximized:

$$(6) \quad L = \sum_{i=1}^n \int_0^{Q^i} VMP_s^i d\beta - \int_0^Z C'(\theta) d\theta - \lambda \left(Z - \sum_{i=1}^n Q^i \right)$$

with respect to the decision variables Q^i and Z , where λ is the usual Lagrange multiplier. The first-order conditions are as follows:

$$(7a) \quad VMP_s^i = \lambda, \text{ and}$$

$$(7b) \quad C'(Z) = \lambda,$$

giving

$$(7c) \quad VMP_s^i = C'(Z),$$

where λ represents the shadow price of water at the source. Equations (7) equate the shadow price λ to the value of marginal product at source for each farmer and to the long-run marginal cost at optimal system capacity. They give the equilibrium condition for optimal allocation of water under spatial efficiency.

Condition (7b) gives Z^* , the optimal stock of water at source.⁵ Let us denote the optimal values of q and Q (those satisfying condition 7) by q^* and Q^* , respectively. Equation (7a) can be rewritten as $\lambda = Pf'(q^*)h(y)$ and totally differentiated to yield

$$(8) \quad dq^*/dy = -f'(q^*)h'(y)/f''(q^*)h(y) < 0,$$

which implies that farmers further from the source receive less (or net) water. Differentiating (2) and using (8) yields $dVMP_r(q^*)/dy = Pf''(q^*)dq^*/dy > 0$. The above discussion can be summarized as follows:

PROPOSITION 1. *Optimal allocation implies that the value of marginal product of water at the source is equal across farmers. The on-farm value of marginal product of water is unequal across farmers and rises with distance from the source. At the optimum the value of marginal product at source equals the total long-run marginal cost of water.*

⁵ Z^* is optimal given an exogenously fixed conveyance system. If the conveyance was endogenously chosen, the optimal water stock might be different.

Figure 2 shows the determination of the shadow price, λ , at the intersection of the aggregate marginal benefit and marginal cost curves. The aggregation of marginal benefit curves is in units of source water. There is a unique transformation from units of source water to units of water received for each farmer through condition (1). Here, λ is equated to individual VMP_r curves to give the optimal source water allocation $Q^*(0)$ at $y = 0$, and $Q^*(y)$ at any downstream location y . Water received by a farmer at the source $q^*(0)$ is the same as water sent, while an amount $q^*(y) = Q^*(y)h(y)$ is received at any downstream location y after conveyance losses.

Spatial Allocation of Source Water

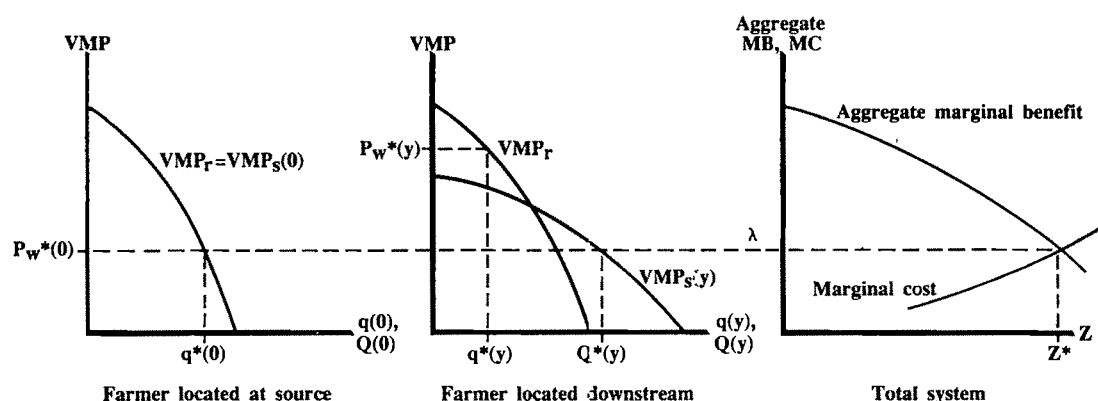
The optimal allocation of source water for each farmer can now be determined by differentiating (1) and substituting (8) to get

$$\begin{aligned} dQ^*/dy &= [-f'(q^*)h'(y)/f''(q^*) - q^*h'(y)]/h^2(y), \\ &= [1/\varepsilon - 1] h'q^*/h^2, \end{aligned}$$

where $\varepsilon = -f''(q^*)q^*/f'(q^*)$ is the absolute value of the elasticity of marginal product of received water in crop production. Hence, $dQ^*/dy = 0$ if either $q^* = 0$ (which is trivial) or $\varepsilon = 1$. Evaluating d^2Q^*/dy^2 at $\varepsilon = 1$ and cancelling terms yields

$$\left. \frac{d^2Q^*}{dy^2} \right|_{\varepsilon=1} = - (d\varepsilon/dy) q^* \cdot h' / h^2 < 0$$

because $d\varepsilon/dy = -f''/f' \cdot dq^*/dy = -f''/$



Note: VMP = value of marginal product, $q(y)$ = water received and $Q(y)$ = source water related by the equation $q(y) = Q(y)h(y)$. MB = marginal benefit, MC = marginal cost, Z = total stock of water, P_w = effective price of water, λ = shadow price of water.

Figure 2. Determination of optimal stock of water, allocation, and pricing

$f' \cdot (-f'h'/f''h) = h'(y)/h(y) < 0$. Therefore, the $Q^*(y)$ function has a local maximum at $\varepsilon = 1$. These results can be combined as follows:

PROPOSITION 2. *Spatial efficiency implies that the allocation of source water increases with distance from source until the absolute value of elasticity of the marginal product for received water equals unity. Optimal water sent decreases with distance beyond that point.*⁶

Note that the length of the canal L (and hence the irrigated area, because we have assumed constant width) is endogenously determined at the point where farmers do not receive any water allocation, or $q^*(L) = 0$.

Spatial Distribution of Charges and Rents

In this section we derive optimal water prices and spot rents accruing to farmers at each location along the canal. The localized shadow price of received water at a distance y is $P_w(y)$. Efficient allocation requires that the marginal product of received water be equated to its shadow price, or $P_w(y) = VMP_r(q^*) = \lambda/h(y)$, using (3) and (7a). At the source, the localized shadow price of water is equal to its system shadow price, since $h(0) = 1$. In addition, $dP_w/dy = -\lambda h'(y)/h^2 > 0$, i.e., the localized shadow price of water increases away from the source.

Marginal cost pricing implies that farmers at any location y pay either the system shadow price of water per unit of source water or the localized shadow price per unit of received water. Both satisfy $\lambda Q^* = (\lambda/h(y))q^* = P_w q^*$ from (1). Intuitively, both marginal cost pricing solutions are equivalent because the total products from source water and received water are equal [from equation (7), marginal cost pricing is the same as marginal product pricing]. It can be shown, by changing the variable of integration, that the areas under the corresponding value of marginal product curves for source water and received water are equal, or

$$(9) \quad \int_0^{q^*} VMP_r d\alpha = \int_0^{Q^*} VMP_s d\tau,$$

where α and τ are variables of integration.

⁶ In the presence of topographical constraints, q^* might lie within the elastic segment of the VMP_r function, in which case $Q^*(y)$ will be a monotonically increasing function of y without an interior maximum.

At marginal cost prices, spot rents from land $R(y)$ at location y are given by

$$(10) \quad R(y) = Pf(q^*) - P_w q^*.$$

Differentiating (10) with respect to y yields

$$R'(y) = dq^*/dy[Pf'(q^*) - P_w] - q^* \cdot dP_w/dy < 0$$

because $Pf'(q^*) = P_w$; or spot rents decrease away from the source when each farmer is making the optimal decision given his (her) location.

In order to examine the effect of conveyance efficiency on optimal water prices and rents, (10) is differentiated with respect to $h(y)$ to obtain $dR/dh = q^* \lambda / h^2 > 0$, which indicates that a higher conveyance efficiency increases land rents. Similarly, $dP_w/dh = -\lambda / h^2 < 0$ and $d^2 P_w / dh^2 = 2\lambda / h^3 > 0$, or increased conveyance efficiency lowers the localized shadow price at an increasing rate. We can thus state the following:

PROPOSITION 3. *Under marginal cost pricing, the localized shadow price of water increases with y and decreases with improved conveyance efficiency. Land rents net of water charges decrease with y and increase with improved conveyance efficiency.*

These results explain the behavior of farmers in response to any given shadow price of source water λ . All farmers pay the same shadow price per unit of source water. As distance from the source increases, the divergence of VMP_r from VMP_s increases, so that the effective price per unit of received water becomes increasingly greater than the price per unit of source water.

The marginal cost pricing solution is shown in figure 2. At the source, the localized shadow price $P_w^*(0) = \lambda$, corresponding to the optimal quantity received $q^*(0)$, is read off the VMP_r curve. For a farmer located at any y , the price for $Q^*(y)$ units of source water is λ . The localized shadow price for $q^*(y)$ units of received water is given by the VMP_r curve, $P_w^*(y)$. Under marginal cost pricing, water charges for the downstream farmer could be expressed as $\lambda Q^*(y)$ or $P_w^*(y)q^*(y)$.

Even though marginal cost pricing is efficient, it may still be objectionable on equity grounds, given that rents near the system head are much higher than those at the tail of the system and because farmers in the middle and tail of the system may pay more money for less total water received.⁷ Moreover, marginal cost pricing

⁷ "Head" and "tail" are relative terms that distinguish the location of the two parties *vis-à-vis* one another.

ing is not necessary on fiscal grounds. Because marginal cost is rising, marginal cost pricing would lead to more than 100% cost recovery. If rationing mechanisms are available to set quantity independently of water charges, then efficient allocation can be achieved without resorting to marginal cost pricing. This facilitates the apparent separability of efficiency and equity and permits a range of possible taxation schemes, including the following:

Rule 1. Proportional benefit taxation. Project beneficiaries could pay in proportion to their individual benefits (Wicksell). This scheme would involve tail farmers receiving less net rents than head farmers, although the benefit distribution would be less skewed than under marginal cost pricing.

Rule 2. Equal rents. The total cost of the system is to be shared among farmer beneficiaries so as to equalize the net rents for all farmers in the system. However, equalizing rents would induce rent-seeking pressures at the boundaries of irrigation systems for system expansion. This is a serious problem of irrigation design in many parts of the world where the target irrigated area is much larger than what the system capacity can efficiently support (Repetto).

Rule 3. Equal charges. Farmers are charged equal amounts for membership in the system. This may provide a pragmatic compromise where the administrative costs and political feasibility of collecting differential water charges is a concern. Equal charges do not achieve equal rents but at least eliminate the "more money for less water" inequity of marginal cost pricing. These alternative financing schemes are illustrated below.

Low versus High Conveyance Losses: An Application

The above model is illustrated by taking conveyance and production functions that closely represent the physical and engineering characteristics of irrigation systems. Representative patterns for the price, quantity, and rent functions are obtained under spatial efficiency for conditions typical to actual irrigation systems.

We construct a *VMP*, function from a survey of functional water-use yield relationships (Hillel) and from linear programming studies of cash crops in Pakistan (Carruthers and Clark). Hillel approximates a yield-water use curve with two segments—a flat portion with constant marginal product and a downward-sloping portion with decreasing marginal product at higher quantities of water:

(11)

$$VMP_r = \begin{cases} 1.75 & \text{for } 0 \leq q \leq 0.3 \\ 2.66 - 3.0 \cdot q & \text{for } 0.3 \leq q \leq 0.875, \end{cases}$$

where *VMP_r* and *q* are in U.S. cents and meters of water, respectively. The two-part function is more accurate than the uniformly downward-sloping curve usually assumed because, at low quantities of water, on-farm losses through deep percolation and evaporation are proportionately higher, justifying a relatively high but constant marginal product over some range of water use. The above function can be treated as a special case of the more general concave, downward-sloping functions analyzed in the previous sections.

The loss of water from seepage and evaporation is assumed proportional to the volume of water carried in the canal (e.g., Tolley and Hastings). Thus, (1) becomes

$$(12) \quad q = Q \cdot e^{-ay},$$

where *a* is the conveyance loss coefficient. Values of *a* from 0.015–0.02 are commonly found in developing country irrigation systems, where canals are often unlined or have linings of low quality, maintenance is poor, and evapotranspiration is high because of higher ambient temperatures (Bos and Nugteren, Hillel). In the simulations that follow, alternative values for *a* of .01 and .02 have been considered.

We assume a hypothetical long-run marginal cost function

$$(13) \quad C'(Z) = Z \cdot 10^{-6}/30,$$

where *C'()* is in cents and *Z* is in cubic meters. Although the marginal cost function for water supply can vary with specific irrigation technologies, projects, and geographical region, a linear form is assumed for simplicity. The total width of the system is 1,000 meters (500 meters on each side of the canal) with a uniform farm size of 5 hectares, which means that canal outlets to successive farms are located every 50 meters.

By taking a starting value of *Z*, we can com-

pute $C'(Z)$ and λ from (13) and (7b). Conditions (7a) and (3) give VMP_s and VMP_r at $y = 0$. From (11), we obtain q^* and Q^* . Because a is known, $h(y)$ is known; and, hence, $P_w(y)$, $q^*(y)$, $Q^*(y)$, and $R(y)$ can be directly obtained. This process is continued for each y until the total stock of water Z is exhausted. By iterating on Z , we choose the optimal Z^* that maximizes total net benefits as defined by the objective function in (5). The efficient length of the canal L is thus determined endogenously.

Simulation Results

Figure 3 and table 1 show the water allocations and charges under alternative conveyance losses and taxation schemes. The results can be summarized as follows:

(a) If conveyance efficiency is low, the aggregate marginal benefit curve shifts up, leading to increased aggregate water use, a higher shadow price of water at source, and a larger irrigated area (see fig. 3, where the efficient length of the system increases from 37.2 kms. to 54.5 kms.).⁸ In this case, reducing the conveyance loss coefficient by one-half leads to expansion of the command area by almost 50%.

(b) Under both high and low conveyance loss regimes, source water allocations increase away from the head and are maximized at about 25 kilometers from the source (see fig. 3). Farmers

⁸ Because width is assumed constant, irrigated area is proportional to canal length.

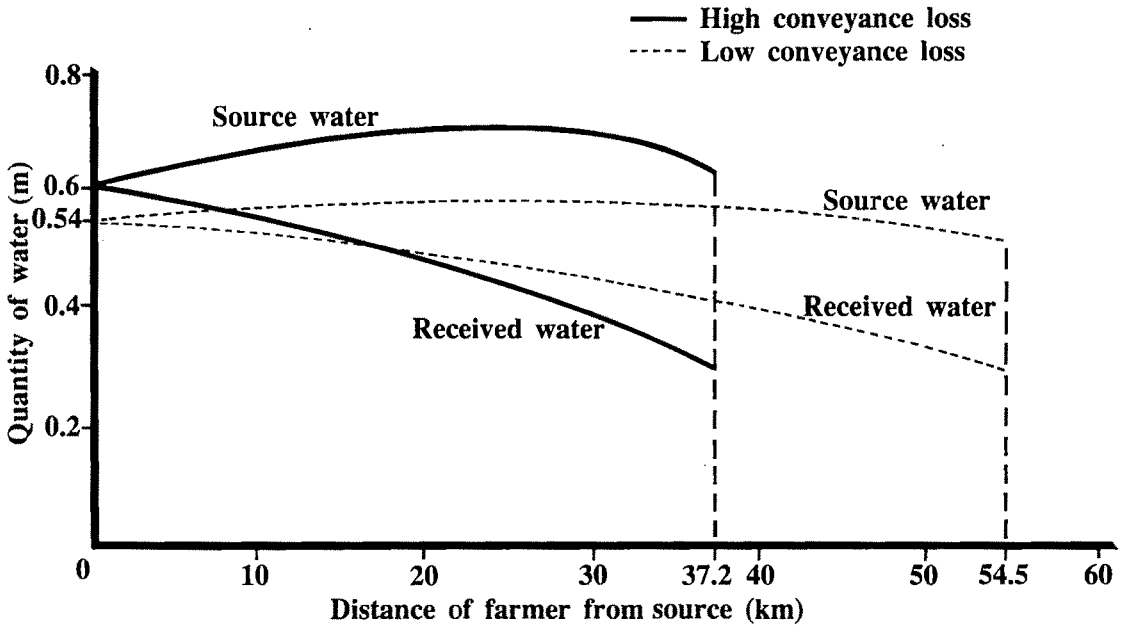


Figure 3. Spatial variation of source water and received water for Pakistan

Table 1. Effect of Alternative Taxation Schemes on Water Charges and Rents for Pakistan (with high conveyance loss)

	Water Charges per Farm (\$)			Rents per Farm (\$)		
	Head	Qmax ^a	L ^b	Head	Qmax	L
Marginal cost pricing	250.1	291.0	262.5	207.2	68.9	0
Proportional benefit taxation	247.0	194.3	141.8	210.4	165.5	120.8
Equal charges	207.8	207.8	207.8	249.6	152.1	54.8
Equal rents	282.2	183.6	86.6	175.0	175.0	175.0

^a Qmax denotes the location where source water allocations are maximized.

^b L denotes the system boundary.

beyond that point get decreasing amounts of source water.⁹

(c) Received water decreases steadily with distance from source (fig. 3). With an increase in conveyance efficiency, the curve for received water flattens, reducing spatial differences in water received. Under high conveyance losses, efficient allocations at the head are approximately double that at the tail.

(d) If farm size is uniform, total water charges under marginal cost pricing are maximized at the point of maximum source water and do not vary substantially over space (table 1).¹⁰ Thus, schemes that charge spatially uniform taxes but ration efficient water allocations (q^*) might be an attractive policy option.

(e) When compared to the other taxation schemes discussed, marginal cost pricing is most inequitable, while proportional benefit taxation permits all farmers to collect a higher degree of rents than if there was no taxation (table 1). The scheme that equalizes net rents across locations benefits farmers located at the tail of the system the most while equalizing water charges provides administrative simplicity.¹¹

Concluding Remarks

In the presence of conveyance losses, efficient spatial allocation of irrigation water implies that the quantity of water applied should fall with distance from the water source. In terms of water produced at the source (before subtracting conveyance losses), the optimal quantity of water allocated to further reaches of the system increases and then decreases with distance from the source.

Under marginal cost pricing, the effective price of water on the farm increases and quasi-rents decrease away from the source. If nonprice instruments for enforcing water allocations are available, then a variety of benefit taxation schemes can be adopted to make the spatial dis-

tribution of rents more uniform, the choice of which depends upon the welfare objectives of the irrigation agency. Water charges that equalize net benefits are vertically equitable but could lead to political pressures for expansion of irrigated area beyond its efficient size. Water charges proportional to net benefits allow higher rents towards the tail of the system, relative to marginal cost pricing, and also preserve horizontal equity and mitigate against rent seeking. Equalizing water charges provides a pragmatic compromise by limiting rents from expansion of the project area, avoiding the additional administrative problems of differential pricing and responding to unfairness associated with paying higher water charges for less total water received.

The model also provides an operational basis for estimating irrigation benefits under alternative allocation rules as the aggregate area under the individual demand curves. Existing methods of estimating irrigation project benefits have largely been based on ad hoc estimates of expected increases in revenues or profits, not on theoretical foundations that take into account the specific characteristics of irrigation systems. The model can also be used to address the concerns of irrigation planners regarding the distribution of benefits and costs in various combinations of allocation rules and financing arrangements.

The benefits of efficient spatial management as described here should be weighed against the higher administrative costs that may be incurred. The model can also be expanded by allowing for the endogenous choice of conveyance and on-farm irrigation technology, and private groundwater extraction. Methods for classifying land quality can also be incorporated (e.g., Caswell and Zilberman). Empirical work is needed to determine the shape of the marginal product schedule under varying environmental conditions. In empirical applications, the framework suggested above can also be extended to estimate the expected benefits of irrigation in different (stochastic) rainfall environments.

[Received March 1988; final revision received March 1990.]

References

- Bishop, A. A., and A. K. Long. "Irrigation Water Delivery for Equity Between Users." *J. Irrigation and Drainage Eng.* 109(1983):349-56.
- Bos, M. G., and J. Nugteren. *Irrigation Efficiency in Small-*

⁹ We performed a similar set of computations with data from California and found that a higher marginal product per unit of on-farm water and higher intensity of water use in California contributed to a larger optimal water stock Z^* and higher effective prices. Similarly, the absolute value of rents, as well as head-tail differentials in prices and rents were much larger in magnitude in California than in Pakistan.

¹⁰ In a low conveyance loss regime, the spatial variation in water charges would be still lower, strengthening the case for uniform water charges over space.

¹¹ Under very high conveyance losses, equalizing net rents could result in tail farmers being subsidized for consumption of irrigation water.

- Farm Areas*. International Commission on Irrigation and Drainage, 1974.
- Bowen, R. L., and R. A. Young. "Appraising Alternatives for Allocating and Cost Recovery for Irrigation Water in Egypt." *Agr. Econ.* 1(1986):35-52.
- Burness, H. S., and J. P. Quirk. "Appropriative Water Rights and the Efficient Allocation of Resources." *Amer. Econ. Rev.* 69(1979):25-37.
- Carruthers, I., and C. Clark. *The Economics of Irrigation*. Liverpool, U.K.: Liverpool University Press, 1981.
- Caswell, M., and D. Zilberman. "The Effects of Well Depth and Land Quality on the Choice of Irrigation Technology." *Amer. J. Agr. Econ.* 68(1986):798-811.
- Chakravorty, U., E. Hochman, and D. Zilberman. "Resource Allocation, Pricing, and Rents for an Impure Public Good: The Case of Irrigation." University of California, Berkeley, 1989.
- Chaudhry, M. A. "Economics of Alternative Irrigation Water Allocation and Pricing Rules in Pakistan." Ph.D. thesis, Colorado State University, 1985.
- Hillel, D. *The Efficient Use of Water in Irrigation*. World Bank Tech. Pap. No. 64, Washington DC: 1987.
- Koenig, L. "The Cost of Water Treatment by Coagulation, Sedimentation and Rapid Sand Filtration." Report prepared for the U.S. Public Health Service, Division of Water Supply and Pollution Control, Washington DC, 1966.
- Loehman, E., and A. Whinston. "A New Theory of Pricing and Decision Making for Public Investment." *Bell J. Econ. and Manage. Sci.* 2(1971):606-25.
- O'Mara, G. T. *Issues in the Efficient Use of Surface and Groundwater in Irrigation*. Washington DC: World Bank Staff Work. Pap. No. 707, 1984.
- Reidinger, R. "Canal Irrigation and Institutions in North India." Ph.D. thesis, Duke University, 1971.
- Repetto, R. *Skimming the Water: Rent-seeking and the Performance of Public Irrigation Systems*. Washington DC: World Resources Institute Res. Rep. No. 4, 1986.
- Tolley, G. S., and V. S. Hastings. "Optimal Water Allocation: The North Platte River." *Quart. J. Econ.* 74(1960):279-95.
- Wade, R. "The System of Administrative and Political Corruption: Canal Irrigation in South India." *J. Develop. Stud.* 18(1982):287-327.
- Wicksell, K. "A New Principle of Just Taxation." *Classics in the Theory of Public Finance*, ed. R. Musgrave and A. Peacock. London: Macmillan & Co., 1950.
- Yoo, K. H., and J. R. Busch. "Least-Cost Planning of Irrigation Systems." *J. Irrigation and Drainage Eng.* 111(1985):352-68.

Optimal Control of Fish Growth

Oscar J. Cacho, Henry Kinnucan, and Upton Hatch

Fish culture is an efficient means of protein production. However, in contrast to land animals, fish cannot be fed *ad libitum* because feed not consumed within a few hours is decomposed. This paper presents a bioeconomic model for determining cost-effective feeding regimes for pond-reared fish. The interplay among feed allowance, diet quality, and harvest date is explored. The optimal control model used is flexible, incorporates the effects of water temperature on fish appetite, and can be used to gain insight into efficient management of aquacultural production systems in different geographical regions.

Key words: aquaculture, bioeconomics, dynamic optimization, feed efficiency.

Aquaculture is rapidly growing throughout the world. In 1985 it accounted for over 14.5% of total fish landings, an increase of 250% over the preceding decade. In the United States, aquaculture is the fastest growing segment of agriculture, increasing at a rate of over 20% per annum from 1980 to 1987 (Manzi). Fueled by rapidly rising retail prices for fish, low feed cost, production efficiencies made possible by improved understanding of fish biology, and growing consumer preferences for fish, U.S. catfish production has more than quadrupled since 1980 (U.S. Department of Agriculture). Mississippi is the largest producer of farm-raised catfish (with over 75% of the total domestic production), followed by Arkansas, Alabama, and Louisiana (Keenum and Waldrop, U.S. Department of Agriculture).

Fish production has several advantages over land animal production systems. Fish convert feed into body tissue more efficiently than land animals, the ratio of weight gain to feed consumed averages 0.84 for channel catfish, compared to 0.48 for broilers and 0.13 for beef. The proportion of edible lean tissue in the dressed carcass of catfish is .81 compared to .60, .54, and .65 for beef, pork, and chicken, respectively

(Lovell). However, the fact that fish are raised in water poses complications not present in land animal production systems. Loss of nutrients occurs if feed is not consumed immediately. After a few hours, unconsumed feed is decomposed by bacteria. The excess feed represents a wasted resource which, when coupled with the added organic wastes discharged by fish, decreases available dissolved oxygen in the water. Low dissolved oxygen decreases fish appetite and growth rates and increases potential death losses. Reduced water quality also exacerbates off-flavor problems (Kinnucan et al.). Although mechanical aeration can mitigate the effects of reduced dissolved oxygen, production costs are increased.

An important decision faced by fish producers is how much to feed each day (Jensen). The distinction between *ad libitum* and satiation feeding is important. In *ad libitum* feeding, feed is available throughout the day, and the animal decides when and how much to consume. In satiation feeding, as much feed as fish will consume in fifteen to thirty minutes is offered once a day. To avoid the risk of accidental overfeeding, Jensen recommends adjusting the feed allowance every two weeks, and feeding approximately 90% of satiation. Thus, reliable estimates of fish appetite under different environmental conditions are needed.

The advantages of using biological and nutritional information in dynamic production models were discussed by Fawcett and later exploited by Chavas, Kliebenstein, and Crenshaw for swine production; Chavas and Klemme for dairy farms; and Talpaz et al. for broilers. Concerning aqua-

The authors are, respectively, a postdoctoral fellow and associate professors, Department of Agricultural Economics and Rural Sociology and Alabama Agricultural Experiment Station, Auburn University.

This is Scientific Journal Paper No. 1-902775P of the Alabama Agricultural Experiment Station.

The authors express appreciation to Tom Lovell and two anonymous reviewers for helpful comments on an earlier draft. Responsibility for the final content rests with the authors.

culture, emphasis has focused on the stocking/harvesting decisions for shrimp (Karp, Sadeh, and Griffin), prawn (Leung and Shang), and carp (Talpaz and Tsur). These studies, however, assume that feeding follows a predetermined schedule. The purpose of the research reported in this paper is to develop a model for determining cost-effective feeding regimes for pond-reared fish and to quantify the interplay among feed allowance (ration size), diet quality (protein percent), and harvest date. The research procedure employs an optimal control model based on a fish growth simulator. Catfish is used because of its importance and because experimental data are available to calibrate the growth model.

This paper provides the basis for applying an iterative research process to aquacultural products. Existing research information is integrated to provide insight into further research needs. Results from new research can in turn be incorporated into the model with no need for major modifications of the basic equations that describe the growth of the animal (Allen et al.). A firm-level optimization problem with dynamic and static components is developed. The biophysical model is briefly discussed, and an optimal control model is developed to solve the dynamic portion of the problem. The strategy used in this paper is similar to that of Talpaz et al. for broilers, in that the economically optimum body-weight trajectory is obtained by controlling feed intake.

The Model

In a typical catfish operation, fingerlings measuring 4 to 6 inches are stocked in spring at 2,500 to 5,000 fish per acre. Feed containing 32% protein is offered every day, either to satiation or at a rate of 3% of estimated body weight per day. Water quality, as measured by the concentration of dissolved oxygen (DO), is checked on a daily basis during summer. When DO falls below a critical level (generally 2 ppm.), aeration equipment is used to restore the oxygen to a safe level. Ponds are generally harvested by processors, who prefer fish weighing 0.75 to 1.25 pounds (340 to 568 grams).

The model developed here is concerned with short-run decisions. Therefore, only variable costs are considered. Production costs and returns are expressed on a per-hectare (ha) basis. The producer seeks to maximize return (π) above vari-

able costs per hectare of pond water. Maximize

$$(1) \quad \pi = YP_Y - (C_F + C_A + C_L),$$

where Y is the quantity of fish (kg/ha), P_Y is price of catfish (\$/kg), C_F is feed cost (\$/ha), C_A is aeration cost (\$/ha), C_L is total fingerling cost (\$/ha). The production function is

$$(2) \quad Y = W_H S,$$

where S is stocking rate (fish/ha), and W_H is harvest weight (kg/fish), defined as

$$(3) \quad W_H = W_0 + \int_{t_0}^{t_H} \omega_t dt,$$

where t_0 is stocking day, t_H is harvest day, W_0 is stocking weight (kg/fish), and ω_t is the growth rate on day t (kg/day).

Production costs are defined by

$$(4) \quad C_F = P_F S \int_{t_0}^{t_H} R(t) F_t dt, \quad 0 \leq R(t) \leq 1,$$

$$(5) \quad C_A = P_K \int_{t_0}^{t_H} A_t dt, \text{ and}$$

$$(6) \quad C_L = P_L S,$$

where P_F is price of feed (\$/kg), P_K is price of electric power (\$/kW · hr), P_L is price of fingerlings (\$/fish), R is ration size (relative to appetite), F_t is fish appetite on day t (kg/fish day), A_t is energy used for aeration (kW · hr/ha day).

In addition we have the following technical relationships:

$$(7) \quad F_t = \frac{dF}{dt} = f\{W(t), T(t)\},$$

$$(8) \quad \omega_t = \frac{dW}{dt} = g\{W(t), T(t), R(t), D\},$$

$$(9) \quad A_t = \frac{dA}{dt} = h\{W(t), T(t), R(t), D, S\},$$

where W is fish weight (kg), T is water temperature (°C), and D is diet composition (percent protein). Functions (7) and (8) represent daily input and output, respectively, simulated by a system of fifteen nonlinear differential equations (Cacho). Simulation starts with food intake (7), and follows the energy flow inside an individual fish to estimate growth rate (8). A simplified version of the model is presented in the following section.

The Biophysical Model

In the fish growth model, the energy provided by feed is partitioned into maintenance and growth needs. Daily growth is treated as a residual quantity after maintenance needs are met (Chavas, Kliebenstein, and Crenshaw). Daily growth rate is defined as

$$(10) \quad \omega_t = \phi(DE R F_t - M_t).$$

The term in parentheses represents daily energy gain, and the parameter ϕ is a function of fish body composition that converts energy into weight units. The value of ϕ ranges from 0.33 to 0.86, depending on the proportion of water, protein, and fat in the fish body; DE is the digestible energy content of the ration (fixed at 2,800 Kcal/Kg); M_t is the energy required for maintenance (Kcal/day); and the remaining variables are as previously defined.

Feed intake is the most important factor influencing fish growth. Fish appetite on a given day is defined as

$$(11) \quad F_t = \Omega_T W^{0.7},$$

where

$$(12) \quad \Omega_T = \frac{\Omega_1 T e^{-\Omega_2/T}}{1 + \Omega_3 e^{-\Omega_4/T}}.$$

All the parameters in equations (11) and (12) are positive. Because fish are cold-blooded, water temperature is a major factor affecting all levels of fish growth. Temperature affects early growth, food intake, maintenance requirements, metabolic rates, and the composition of gain (Kau-shik). Ω_T increases from close to zero at 8 degrees centigrade ($^{\circ}\text{C}$) to a maximum of 0.43 at 33°C and rapidly decreases thereafter.

Energy used for maintenance is obtained from dietary protein, fat, and carbohydrates. The latter nutrients are readily converted to fat and stored in that form. Thus maintenance energy (M_t) is defined as

$$(13) \quad M_t = MP_t + MF_t,$$

where MP_t and MF_t are maintenance energy provided by protein and fat, respectively, defined as

$$(14) \quad MP_t = \beta_P R^{\tau_P} \Gamma_T W^{0.8},$$

$$(15) \quad MF_t = \beta_F R \Gamma_T W_K^{0.8},$$

where

$$(16) \quad \Gamma_T = \Gamma_1 e^{\Gamma_2 T}.$$

In (15), W_k represents fish body energy. All the parameters in equations (14) to (16) are positive, and τ_P is greater than 1. β_P , τ_P and β_F are functions of diet composition (digestible protein/digestible energy).

Substituting equations (11) to (16) into (10) yields a growth function that is concave in both ration size and water temperature (with a maximum when $T = 30^{\circ}\text{C}$). This model simulates protein and fat dynamics, allowing for the inclusion of quality attributes into the production function. Coefficient values and other details of the growth model are presented by Cacho.

The expected annual water temperature cycle is described by the function

$$(17) \quad T(t) = T_A + T_R \sin \left(2\pi \frac{t_0 + t - t_A}{365} \right),$$

$$1 \leq t_0 \leq 365,$$

where T_A is the average annual temperature, T_R represents the range of temperature about T_A , t_A is the time of the year at which $T = T_A$, stocking date (t_0) equals 1 on January first, and t is the number of days from stocking to the current day. It was assumed that $T_A = 20^{\circ}\text{C}$, $t_A = 120$ days, and $T_R = 13^{\circ}\text{C}$. This function was added to the model to provide the average daily temperature input; it can be applied to different geographical areas by specifying the appropriate parameters.

Energy used for aeration [equation (9)] is estimated by

$$(18) \quad A_t = \frac{DO_d}{OT_p},$$

where DO_d is oxygen demand (Kg/ha day) and OT_p represents oxygen transfer rate in the pond (kg/kW hr). The oxygen requirements of a fish are given by diet composition and energy used for maintenance. These requirements were estimated from M_t in equation (13), by using oxycaloric equivalents of specific nutrients (Elliott and Davidson). OT_p is defined as (Boyd, Rajendren, and Durda)

$$(19) \quad OT_p = OT_s \frac{\beta DO_s - DO_p}{9.07} 1.024^{T-20} \alpha,$$

where OT_s is standard oxygen transfer rate (kg/kW hr) (provided by the aerator manufacturer); DO_s is oxygen saturation level at actual water temperature, obtained from solubility tables (Boyd); DO_p is the dissolved oxygen level to be maintained in the pond; $\alpha = 0.77$ and $\beta = 0.94$ are correction factors determined by Shelton and

Boyd for Alabama ponds. A value $OT_s = 1.9$ is close to the average for electric paddlewheel aerators (Ahmad and Boyd). $DO_p = 5$ mg/l is assumed because this level will provide enough oxygen without constraining appetite or growth.

The following assumptions were made in solving the optimization problem: (a) Temperature is deterministically given by time of the year as described by equation (17). (b) Feed is purchased in bulk; therefore, diet composition is fixed throughout the growing season. (c) Six thousand 20-gram fish/hectare are stocked on April first. (d) The rate of aeration (A_i) is automatically adjusted in response to the fish oxygen demand to keep the dissolved oxygen level at 5ppm.

The stocking rate assumption deserves comment. Although increasing the stocking density increases total production per unit area (up to the carrying capacity of the pond), the higher stocking increases production risks (Hanson, Martin, and Flynn) and produces greater variability in fish size at harvest. The effects of stocking density on individual fish growth are complex because of the interplay of social interactions (e.g., competition for food and aggression) and the effects of higher biomass on water quality.

Assumption (c), when coupled with assumption (d), represents a conservative approach. The pond environment is maintained as favorable as possible, reducing the risk of losses due to low dissolved oxygen and diseases. These assumptions can be relaxed once more experimental information becomes available to allow their treatment as decision variables in the growth model.

Solution Strategy

By substituting equations (2) through (19) into (1) we can express profit in terms of nine variables. The state variables $W(t)$ and $DO_p(t)$ (endogenous) and $T(t)$ (exogenous) describe the state of the system at any time t . The decision variables t_0 , t_H , $R(t)$, A_i , D , and S can be controlled. $DO_p(t)$ is fixed at 5ppm through adjustment of the decision A_i [assumption (d)]. $T(t)$ cannot be controlled; its trajectory from t_0 to t_H is predetermined by assumption (a). Because S and t_0 are fixed [assumption (c)], the problem is reduced to finding the optimal combination of $R(t)$, D , and t_H . Because fish are fed once a day, time is divided into discrete one-day intervals. We obtain a set of $n = 2 + (t_H - t_0)$ first-order conditions of the form,

$$(20) \quad \frac{\partial \pi}{\partial x_i} = P_F \frac{\partial Y}{\partial W_H} \frac{\partial W_H}{\partial x_i} - \left(\frac{\partial C_F}{\partial x_i} + \frac{\partial C_A}{\partial x_i} \right) = 0, \\ \text{for } i = 1, 2, \dots, n;$$

where $x_1 = D$, $x_2 = t_H$, $x_3 = R(t_0)$, ..., $x_n = R(t_H)$. The complexity of the biophysical model and the presence of nonlinear equations preclude an analytical solution to the problem.

Numerical solution was accomplished in two steps. In the first step, an optimal control algorithm was used to optimize the dynamic decision variable $R(t)$ for different values of the static decision variables D and t_H , and for different terminal values of the state variable $W(t_H)$. Optimal control results were output to a file that, in the second step, was used to find the profit-maximizing solution. The second step can be easily accomplished using a spreadsheet or graphics package.

The Optimal Control Problem

The problem is to find the trajectory of the control variable $R(t)$ that minimizes the cost of producing a fish of weight W_H at harvest time t_H . The objective is to minimize

$$(21) \quad J = P_F \int_{t_0}^{t_H} R(t) F_t dt,$$

subject to

$$(22) \quad F_t = f\{W(t), t\},$$

$$(23) \quad \omega_t = g\{W(t), R(t), t\},$$

$$(24) \quad 0 \leq R(t) \leq 1,$$

$$(25) \quad W(t_0) = W_0, \text{ and}$$

$$(26) \quad W(t_H) = W_H.$$

The functional (J) represents the cost of feeding an individual fish from stocking to harvest. Because P_F is constant for a given diet composition (D), the problem can be reduced to maximize

$$(27) \quad H = -R(t) F_t + \sigma(t) W(t);$$

constraint (24) is introduced by the lagrangean function:

$$(28) \quad \Lambda = H + \lambda_1(t)R(t) + \lambda_2(t)[1 - R(t)],$$

which translates into the optimal control function,

(29)

$$R^*(t) = \begin{cases} 0 \\ R^S(t) \in [0, 1] \\ 1 \end{cases} \quad \text{if} \quad \begin{cases} H_R < 0 \\ H_R = 0 \\ H_R > 0 \end{cases};$$

where

$$H_r = \frac{\partial H}{\partial R}.$$

Here, H_R is the switching function and R^S is the singular path of the control. Bang-bang control occurs when the singular path is outside the allowed bounds (Beltrami, Liu and Forker).

Evaluating the Euler equation for the state W (Burghes and Graham), we obtain the adjoint equation:

$$(30) \quad \frac{d\sigma}{dt} = R(t) \frac{\partial F_t}{\partial W} - \sigma(t) \frac{\partial \omega_t}{\partial W}.$$

Equation (29) implies that along the singular path:

$$(31) \quad \sigma^*(t) = \frac{F_t}{\frac{\partial \omega_t}{\partial R}}; 0 \leq R \leq 1.$$

The adjoint function shows that the shadow cost on a given day is represented by the ratio of fish

appetite to the rate of change in fish growth rate with respect to the control variable.

F_t places an upper limit on the amount of energy that can enter the system on a given day; F_t is determined by temperature and weight (equation 11) and is not affected by the control $R(t)$. For a given state of the system, an increase in the efficiency of feed use will cause a decrease in the shadow cost.

The functions involved in equation (31) are illustrated in figure 1. This figure was numerically generated using equations (10) to (16). Figure 1A shows fish appetite [the numerator of equation (31) or equation (11)] in temperature space at five different fish weights. Four states (I to IV), corresponding to different temperature/weight combinations, were selected to analyze $\sigma(t)$. The shadow cost increases with increasing R (fig. 1B); roman numerals on the curves indicate the corresponding state of the system. At $T = 30$ (states I and III), the increase is nearly linear. At $T = 33$ (states II and IV), the shadow cost increases at an increasing rate. When the system is at a high- T high- W combination (IV), a rapid increase in shadow cost occurs as the control increases, implying a higher pressure to use feed efficiently.

If the differential equation (30) could be solved

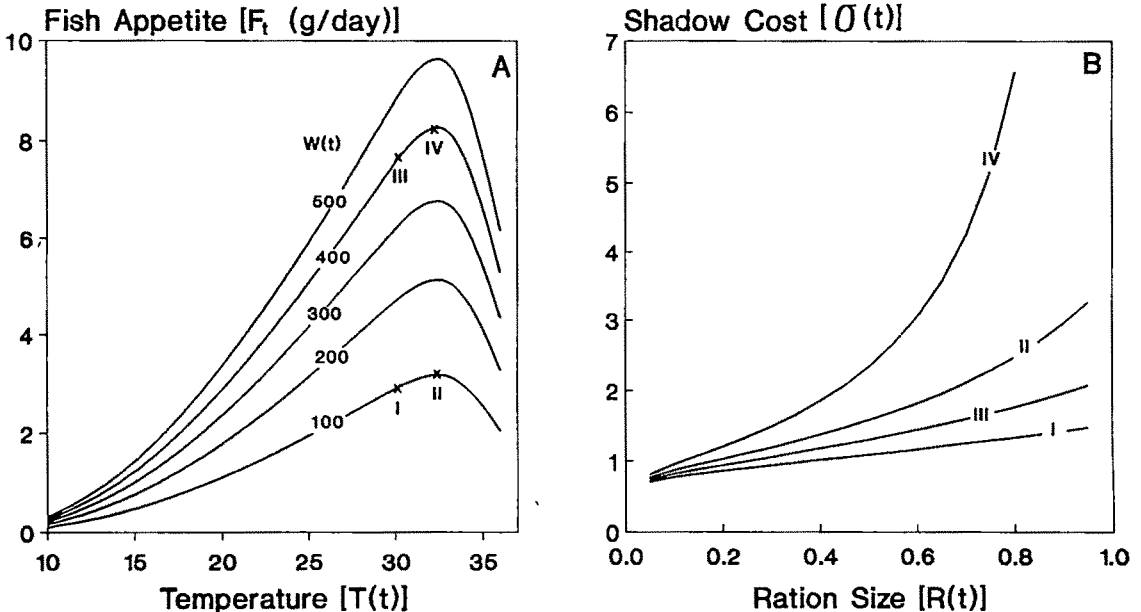


Figure 1. Relationship between the control $R(t)$ and the shadow cost $\sigma(t)$ at four different states of the system (temperature/weight combinations)

for $\sigma(t)$ as a function $R(t)$, the result would be a system of two equations (30) and (31), with two unknowns ($R^*(t)$ and $\sigma^*(t)$). While these equations are useful for gaining insight into the system, numerical solution of the problem is more easily obtained by finding the value of $R(t)$, within the interval $(0, 1)$, that maximizes H in equation (27). Then equation (30) is used to estimate the value of $\sigma(t + 1)$. Solution of equation (30) requires an estimate of the initial value of the adjoint $\sigma(t_0)$. Although $\sigma(t_0)$ is unknown, the final value of the state $W(t_H)$ is known, providing the boundary necessary to solve the system. If the terminal value of the state were open, the transversality condition $\sigma(t_H) = 0$ would apply (Seierstad and Sydsæter). However, in the dynamic version of the traditional cost minimization problem, with $W(t_H)$ fixed, $\sigma(t_H)$ is not constrained.

A shooting algorithm combined with a Brent maximization subroutine (Press et al.) was used for numerical solution. The algorithm consists of estimating an initial value for the adjoint, $\sigma(t_0)$, and numerically integrating equation (30) while maximizing equation (27) at one-day intervals. If the terminal condition in equation (26) is not achieved, $\sigma(t_0)$ is adjusted and the procedure is repeated until $W(t_H) = W_H$ is obtained. At this point the optimal solution for the given set of constraints occurs. For given values of the static decision variables D and t_H , the corresponding functions $R^*(t)$, $\sigma^*(t)$, and $W^*(t)$ represent, respectively, the optimal trajectory for the control, the adjoint, and the state variable.

Results

Optimal Trajectory

Figure 2 shows the optimal trajectories compared to a growth maximization and the feeding schedule recommended by Jensen. A 32% protein diet, the most commonly used by catfish farmers (Keenum and Waldrop), was used to generate these curves. The effects of water temperature on fish appetite and potential for growth are apparent. For the assumed temperature cycle, maximum feed consumed (F_{max}) increases from almost zero to a maximum of 38 kg per ha per day in late summer (fig. 2A). The optimal control path ($R^*(t)F_t$) is a bimodal curve with a local maximum in late spring, a local minimum in midsummer, and a global maximum in late summer. This trajectory minimizes the area un-

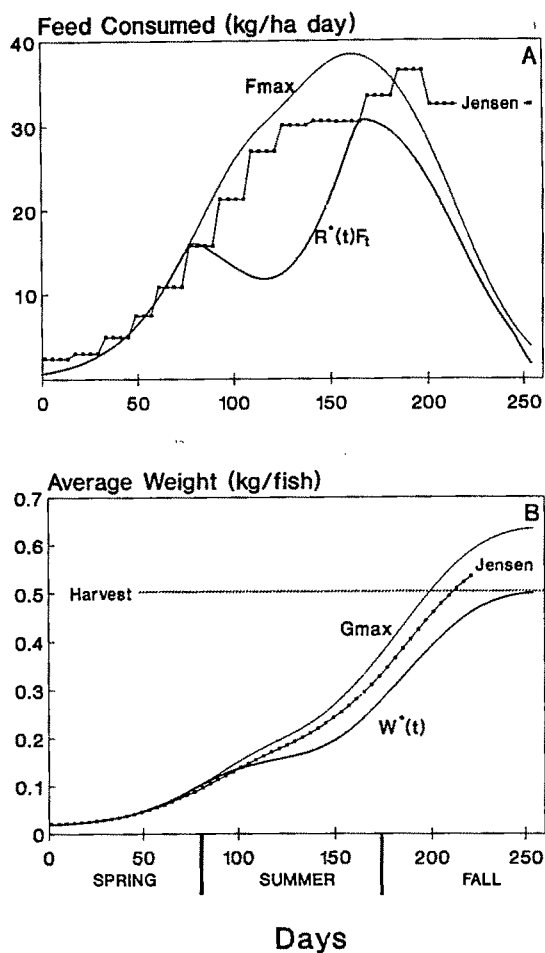


Figure 2. Trajectory of the state variables $T(t)$ and $W(t)$, and the control variable $R(t)F_t$ under optimal control, Jensen's feeding schedule, and maximum growth

der the curve while satisfying the boundary conditions of the problem. The integral of $R^*(t)F_t$ represents total feed necessary to produce a fish of weight W_H on day t_H with a $D\%$ protein diet. The shape of the curve is the same for different values of D , t_H , and W_H ; but the extreme points change. Thus, feed input use is affected by the static decision variables. In contrast to growth maximization, in which fish are fed to satiation, the cost minimization solution implies that feed should be restricted during selected growth periods (see fig. 2A). The restricted feeding regime is consistent with the results for broilers obtained by Talpaz et al. Jensen's stepwise feeding schedule falls between the optimal and maximum trajectories during the summer months.

The fish growth curves (fig. 2B) exhibit three

inflection points. Fish weight increases at an increasing rate during spring; after the first inflection point in late spring, it increases at a decreasing rate in midsummer. The second inflection point occurs in late summer, and fish gain weight at an increasing rate during fall. The third inflection point occurs in late fall, and fish growth rate rapidly decreases as winter approaches. The patterns of growth in the curves shown in figure 1B are caused by the temperature cycle, in concert with the associated feeding trajectory and the changing needs of the growing fish [see equations (11) to (17)]. The optimal path of the state variable, $W^*(t)$, gives the slowest growth rate during summer, followed by Jensen's trajectory. A harvest weight of 500 grams is reached in 198, 210, and 250 days under maximum growth, Jensen's feeding schedule, and optimal control, respectively. Because growth rate is so slow during winter, it is not practical to stock new fish in the fall. Thus, the only benefit of a shorter season is the reduced interest payment.

Isoquant Analysis

The optimal control path was determined for 180 pairs of dietary protein (26% to 42% at increments of 2) and crop length (205 to 250 days at increments of 5) values, at 2 harvest weights (500 g and 600 g). Total feed consumed was estimated by numerically integrating $R^*(t)F_t$. Preliminary analysis indicated that total feed consumed and the initial shadow cost can be approximated by the function:

$$(32) \quad Z = \alpha + \beta_1 C + \beta_2 D + \beta_3 D^2 + \beta_4 CD,$$

where Z represents either $\sigma^*(t_0)$ or total feed consumed (kg/ha), and C represents crop length (days). Results for $\sigma^*(t_0)$ can be incorporated into the optimal control algorithm to provide an initial estimate for a given set of boundary conditions. The regression results for feed consumed can be used to approximate the protein/crop-length pair that minimizes cost, thereby reducing the computational effort involved in finding a global minimum. An iterative procedure can be established, whereby the optimal control algorithm is run at smaller intervals in the neighborhood of the previous approximation.

After analysis of preliminary results, the optimal control solution was obtained for a finer

Table 1. Effects of Crop Length, Dietary Protein, and Harvest Weight on Optimal Control Results

Parameter	$\sigma^*(t_0)$		Feed	
	(harvest weight)			
	500g	600g	500g	600g
α	64.6 (15.71) ^a	375.8 (1.51)	12,089 (40.07)	20,140 (15.09)
β_1	0.027 (1.63)	3.29 (3.26)	-1.34 (1.10)	15.70 (2.91)
β_2	-3.17 (23.58)	-38.43 (4.79)	-411.22 (41.67)	-875.87 (20.37)
β_3	0.055 (51.86)	0.98 (16.04)	5.58 (71.28)	14.91 (45.78)
β_4	-0.0026 (5.52)	-0.12 (4.12)	-0.10 (2.90)	-0.85 (5.47)
R^2	0.98	0.81	0.99	0.98
N	128	122	128	122

Note: estimated coefficients for equation (32).

^a The absolute value of t -ratios are shown in parentheses.

grid of static decision variables, with t_H ranging from 228 to 254 at two-day increments, and D ranging from 31% to 39% at 0.5 increments, and for harvest weights of 500 and 600 grams. Final regression results for equation (32) are presented in table 1. All the coefficients are significantly different from zero ($P < 0.05$), except α for the shadow cost at 600 grams and β_1 for feed at 500 grams. Both variables are convex in dietary protein.

Isoquants were obtained by plotting total feed consumed against the corresponding dietary protein value for a given crop length (fig. 3). Points on these isoquants can be used to find the cost-minimizing level of protein for a given price situation. The actual distribution of feed through the growing season can then be obtained by running the optimal control algorithm. Three points are worth noting: (a) The shape of the isoquants indicates that feed quantity and quality are substitutes in production within a certain range. As the protein content of the diet increases, the amount of feed necessary to obtain a given weight decreases almost linearly between 31% and 34% protein. This is equivalent to Chavas, Kliebenstein, and Crenshaw finding that corn and soybeans are substitutes in swine production. (b) As crop length increases, the amount of feed needed decreases, implying that feed is more efficiently converted into animal tissue by delaying the marketing date. (c) High dietary protein has negative effects on growth of larger fish, as indicated by the isoquants for a 600-gram fish curving upward. This result is also consistent with the findings by Chavas, Kliebenstein, and

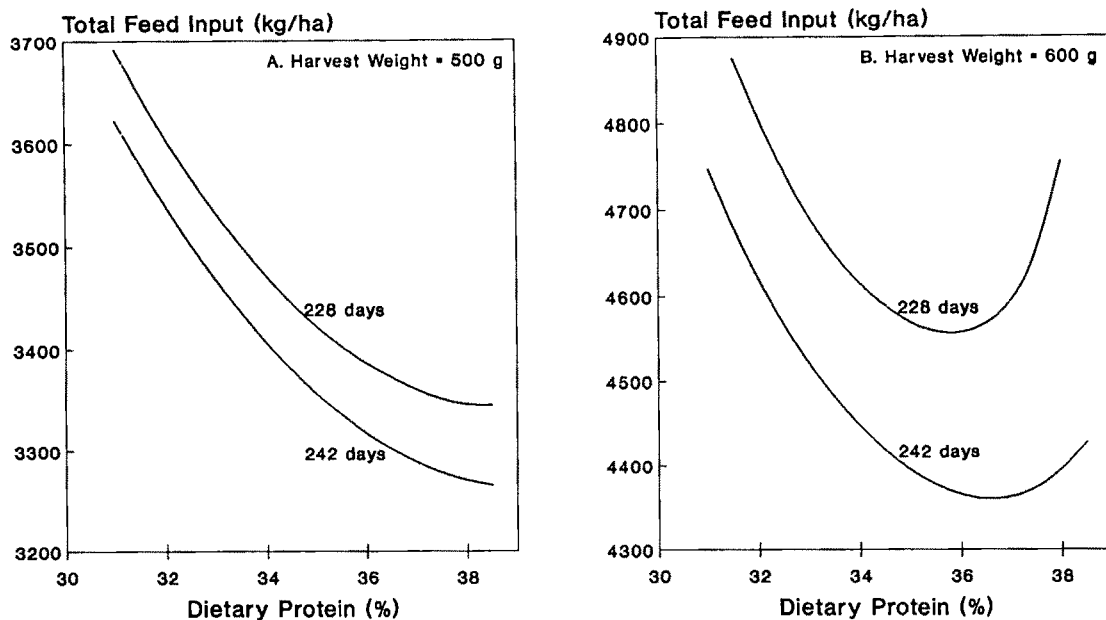


Figure 3. Feed-protein isoquants derived from optimal control results

Crenshaw, where optimal dietary protein decreased as pigs grew larger.

Price Sensitivity

Feed price was determined by the function

$$(33) \quad P_F = \delta_1 + \delta_2 \cdot DP,$$

where P_F is feed price (\$/Kg), and DP is dietary protein content (Kg protein/Kg feed). The intercept represents the nonprotein cost of feed. Its value will be affected by changes in the costs of marketing and processing, as well as by the prices of grains (mainly corn) and fats that provide dietary energy in catfish feeds. The slope of equation (33) represents the protein cost of feed. Its value will be affected by the prices of soybean and fish meal, the two main sources of protein in catfish diets. Both δ_1 and δ_2 can be represented by a probability density function. However, because it is assumed that feed is purchased in one installment at the beginning of the season, these parameters are treated as given.

The coefficients in equation (33) can be estimated from prices of catfish feeds of different protein contents. Alternatively, data can be generated for different diets by a least-cost formulation for a set of ingredient prices. Using the latter method, Reutebuch estimated $\delta_2 = 0.543$; $\delta_1 = 0.112$ was algebraically estimated assum-

ing a cost of \$260 per ton for a 32% protein diet (Crews and Jensen).

Table 2 presents the sensitivity of optimal static decisions to changes in input prices. Dietary protein and crop length that maximize return above variable costs (discounted at 12%) are shown. Optimal protein ranges between 32% and 36.5%. For a given set of prices, optimal protein values are lower for the higher harvest weight. This result suggests the possible benefit of allowing diet composition to become a dynamic decision variable. Interest rate has practically no effect on optimal protein. Energy price has a slight effect on optimal protein, where higher prices are associated with higher protein. This result is not surprising because the demand for energy is related to feed allowance, and higher dietary protein requires lower feeding rates (see fig. 3). Optimal crop length ranges between 232 and 252 days. As expected, crop length is negatively related to interest rate, although the effect is small. Feed and energy prices are positively related to crop length, indicating that they increase pressure to use feed more efficiently (i.e., reduce the amount of feed per unit of weight gain, see fig. 3).

Table 3 shows the ratio of profit obtained with an optimal control trajectory over that obtained with a growth-maximizing strategy (assuming a catfish price of \$1.43/kg or \$0.65/lb). Profit ratios range from 1.02 to 1.27. At base feed

Table 2. Sensitivity of Profit-Maximizing Solution to Input Price Changes

Harvest Weight = 500 grams						
Optimal Dietary Protein (%)			Optimal Crop Length (days)			
Protein price ^a	Non-Protein Price ^b					
	low	base	high	low	base	high
low	35.0	35.5	36.5	238	240	242
base	33.0	34.0	35.0	240	242	244
high	32.0	33.0	34.0	244	244	246
Feed price ^c	Interest Rate ^d					
	low	base	high	low	base	high
low	35.0	35.0	34.5	244	238	232
base	34.0	34.0	34.0	246	242	240
high	34.0	34.0	34.0	248	246	244
Energy Price ^e						
	low	base	high	low	base	high
low	34.0	35.0	35.0	232	238	240
base	33.5	34.0	34.5	242	242	244
high	33.5	34.0	34.0	244	246	246

Harvest Weight = 600 grams						
Optimal Dietary Protein (%)			Optimal Crop Length (days)			
Protein price	Non-Protein Price					
	low	base	high	low	base	high
low	34.0	34.5	35.0	244	246	248
base	33.0	34.0	34.0	246	250	250
high	32.5	33.0	33.5	250	248	252
Feed price	Interest Rate					
	low	base	high	low	base	high
low	34.0	34.0	34.0	248	244	240
base	34.0	34.0	34.0	250	250	246
high	33.5	33.5	33.5	252	252	250
Energy Price						
	low	base	high	low	base	high
low	33.5	34.0	34.5	242	244	246
base	33.5	34.0	34.0	246	250	250
high	33.5	33.5	34.0	250	252	250

Note: Low and high represent the base price times 0.5 and 1.5, respectively.

^a The value of δ_2 in equation (33), base = \$0.543.

^b The value of δ_1 in equation (33), base = \$0.112.

^c A proportional change in A and B.

^d Base = 0.12/year.

^e Base = \$0.055/kW hr.

Table 3. Ratio of Profit Obtained with an Optimal Control Strategy as Compared to a Profit Maximization Strategy

Harvest Weight = 500g Harvest Weight = 600g						
Protein price ^a	Non-Protein Price ^b					
	low	base	high	low	base	high
low ^c	1.06	1.09	1.14	1.02	1.03	1.05
base	1.09	1.14	1.19	1.03	1.05	1.06
high	1.12	1.18	1.27	1.04	1.06	1.09

Note: Assume a catfish price of \$1.43/kilogram.

^a The value of δ_2 in equation (33), base = 0.543.

^b The value of δ_1 in equation (33), base = 0.112.

^c Low and high represent the base price times 0.5 and 1.5, respectively.

prices, profit can be increased by 14% and 5%, respectively, for 500-gram and 600-gram fish by following an optimal trajectory.

Concluding Comments

An optimal control formulation of the decision problem about catfish feeding under aquacultural conditions yields several insights useful for defining more effective production regimes. The key findings are summarized as follows:

(a) In contrast to current industry practices, the optimal control solution suggests fish should be "starved" during periods in which water temperatures are high. Reducing feed intake during the hot summer months improves the efficiency of feed conversion without impairing fish health. Further, restricting the diet helps maintain water quality and may reduce the incidence of off-flavor.

(b) The cost-minimizing feed trajectory differs from the growth-maximizing trajectory, both in terms of total quantity of feed and feeding pattern. At a farm price of \$0.65 per pound profit can be increased by as much as 27% (depending on feed prices) by following an optimal trajectory as compared to growth maximization. The differences suggest significant profit advantages to producers who pursue economic rather than biological optima.

(c) The differences between biological and economic optima lessen as the harvest weight of the fish increases. Thus, the allocative error arising from feeding the fish to satiation (as would be required for growth maximization) may become less important if processors prefer larger fish (over 600 grams liveweight). Currently processors favor 340 to 568-gram fish (Jensen).

(d) Optimal dietary protein ranges from 32% to 36.5%, depending on nutrient prices. The current industry standard is 32% protein (Keenum and Waldrop), and it is not changed in response to price changes. Isoquant analysis suggests that benefits can be attained by allowing dietary protein to change during the growing season.

(e) At base interest rates (12%), the optimum harvest date for 500-gram fish ranges from 238 to 246 days after stocking. Optimum harvest date for a 600-gram fish ranges from 244 to 252 days. As feed price increases, optimal crop length increases, reflecting the pressure to improve feed conversion efficiency by feeding at a reduced rate over a longer time period.

(f) Further development of the model presented here can produce valuable management

aids for fish producers. Additional research is needed concerning the effects of population density and dissolved oxygen on fish appetite and metabolism. Such information would permit stocking density and aeration rate to be treated as decision variables.

[Received October 1989; final revision received March 1990.]

References

- Ahmad, T., and C. E. Boyd. "Design and Performance of Paddle Wheel Aerators." *Aquacultural Eng.* 7(1988): 39-62.
- Allen, P. G., L. W. Botsford, A. M. Schuur, and W. E. Johnston. *Bioeconomics of Aquaculture*. Amsterdam: Elsevier Science Publishers, 1984.
- Beltrami, E. *Mathematics for Dynamic Modeling*. Orlando FL: Academic Press, 1987.
- Boyd, C. E. *Water Quality in Warmwater Fish Ponds*. Auburn University, Alabama Agr. Exp. Sta., 1979.
- Boyd, C. E., R. B. Rajendren, and J. Durda. "Economic Considerations of Fish Pond Aeration." *J. Aqua. Trop.* 1(1986):1-5.
- Burghes, D., and A. Graham. *Introduction to Control Theory Including Optimal Control*. New York: John Wiley & Sons, 1980.
- Cacho, O. J. "Protein and Fat Dynamics in Fish: A bioenergetic Model Applied to Aquaculture." *Ecological Modelling*, 50(1990):33-56.
- Chavas, J.-P., and R. M. Klemme. "Aggregate Milk Supply Response and Investment Behavior on U.S. Dairy Farms." *Amer. J. Agr. Econ.* 68(1986):55-66.
- Chavas, J.-P., J. Kliebenstein, and T. D. Crenshaw. "Modeling Dynamic Agriculture Production Response: The Case of Swine Production." *Amer. J. Agr. Econ.* 67(1985):636-46.
- Crews J., and J. W. Jensen. *Budgeting for Alabama Catfish Production*. Alabama Coop. Extens. Serv., Auburn University, 1987.
- Elliot, J. M., and W. Davidson. "Energy Equivalents of Oxygen Consumption in Animal Energetics." *Oecologia* 19(1975):195-201.
- Fawcett, R. H. "Towards a Dynamic Production Function." *J. Agr. Econ.* 24(1973):543-55.
- Hanson, G. D., N. R. Martin, and J. B. Flynn. "Production, Price and Risk Factors in Channel Catfish Farming." *S. J. Agr. Econ.* 16(1984):173-82.
- Jensen, J. "Channel Catfish Production in Ponds." Alabama Coop. Extens. Serv. Circular ANR-195, Auburn University, undated.
- Karp, L., A. Sadeh, and W. L. Griffin. "Cycles in Agricultural Production: The Case of Aquaculture." *Amer. J. Agr. Econ.* 68(1986):553-61.
- Kaushik, S. J. "Environmental Effects on Feed Utilization." *Fish Physiology and Biochemistry* 2(1986):131-40.
- Keenum M. E., and J. E. Waldrop. "Economic Analysis of Farm-Raised Catfish Production in Mississippi." Mississippi State University Agr. and Forestry Exp. Sta. Tech. Bull. No. 155, 1988.
- Kinnucan, H., S. Sindelar, D. Wineholt, and U. Hatch. "Processor Demand and Price-Markup Functions for Catfish: A Disaggregated Analysis with Implications for the Off-Flavor Problem." *S. J. Agr. Econ.* 20(1988):81-91.
- Leung, P. S., and Y. C. Shang. "Modeling Prawn Production Management Systems: A Dynamic Markov Decision Approach." *Agr. Systems* 29(1989):5-20.
- Liu, D. J., and O. D. Forker. "In Search of Optimal Control Models for Generic Commodity Promotion." *Agr. Econ. Work. Pap.* 88-5, Cornell University, 1988.
- Lovell, R. T. *Nutrition and Feeding of Fish*. New York: Van Nostrand Reinhold Publishers, 1989.
- Manzi, J. "Aquaculture, Research Priorities for the 1990s." *World Aquaculture* 20(1989):29-32.
- Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. *Numerical Recipes. The Art of Scientific Computing*. Cambridge: Cambridge University Press, 1986.
- Reutebuch, E. M. "Quadratic Growth Response of Pond-Raised Channel Catfish to Increasing Levels of Dietary Protein and Its Economic Implications." M.S. thesis, Auburn University, 1988.
- Seierstad, A., and K. Sydsæter. *Optimal Control Theory with Economic Applications*. New York: Elsevier, 1987.
- Shelton, J. L., and C. E. Boyd. "Correction Factors for Calculating Oxygen Transfer Rates of Pond Aerators." *Transactions Amer. Fisheries Soc.* 112(1983):120-22.
- Talpaz, H., S. Hurwitz, J. R. de la Torre, and P. J. H. Sharpe. "Economic Optimization of a Growth Trajectory for Broilers." *Amer. J. Agr. Econ.* 70(1988):382-90.
- Talpaz, H., and J. Tsur. "Optimising Aquaculture Management of a Single-Species Fish Population." *Agr. Systems* 9(1982):127-42.
- U.S. Department of Agriculture. *Catfish*. Washington DC: U.S. Department of Agriculture Crop Reporting Board, Statistical Reporting Service. Various issues 1980-88.

Farm Production Decisions Under Cross and Conservation Compliance

Dana L. Hoag and Herb A. Holloway

The success of soil conservation compliance (CC) hinges on participation in commodity programs. Using mixed integer programming, the profitability of participation in commodity programs was examined on seventeen surveyed North Carolina farms. Without CC, cross compliance in commodity programs increased projected soil erosion by over 80% when participation increased from 30% to nearly 90%. With CC, erosion fell by two-thirds at the high participation level but decreased by only 1% with low participation. Individual farm acreage base and crop yield strongly affected the profitability of CC. Compliance is increasingly profitable on farms with more base acreage and higher yields.

Key words: conservation compliance, cross compliance, North Carolina, soil erosion.

Congress passed several programs in the 1985 Food Security Act to reduce soil erosion on U.S. cropland (Glaser, Ervin). One program, conservation compliance (CC), requires producers to have a plan for applying appropriate conservation measures on highly erodible land. Producers without such a plan are ineligible for most government farm program benefits (e.g., deficiency payments and commodity loans). The conservation plan must be developed by 1990 and fully implemented by 1995. A related program, conservation reserve, offers annual rental payments and cost sharing for removing highly erodible lands from production for ten years.

While CC may be effective in removing commodity program incentives to use erosive farming methods, its impact on soil erosion is limited to program participants. The effectiveness of the conservation reserve is also limited because only farmers in conservation compliance may collect rental payments.

One of the factors influencing the effectiveness of conservation compliance in reducing soil erosion is the costs of implementing a conservation plan relative to commodity program benefits. The net impact of CC on participation costs

will be highly site-specific depending on soil type, climate, historical production, current acreage base, and other farm-level characteristics. These variables and others determine whether CC raises or lowers program participation and, hence, how it affects soil erosion.

Soil erosion may increase where CC leads to lower participation rates in farm programs. Battie and Sappington found that the majority of farmers in a Tennessee watershed would not find it profitable to comply with a hypothetical cross-compliance program. Therefore, erosion might increase if the program were imposed. However, erosion levels could fall if cropping restrictions imposed by commodity programs have led to the production of erosive crops. Producers who no longer participate in farm programs may turn to less erosive crops.

The objective of this analysis is to determine how farm-level variables such as acreage base and average yield affect the impact of conservation compliance on farm profits. Key factors influencing participation were determined by comparing optimal net returns with and without commodity programs on seventeen North Carolina farms. Mixed integer programming (MIP) was used to optimize program participation-crop mix decisions. Perry et al. demonstrated the usefulness of MIP in modeling farm programs but did not examine conservation compliance.

Other researchers have examined the conservation reserve program (Reichelderfer and Boggess, Young and Osborn, Hertel and Preckel) and conservation compliance (Barbarika and

Dana L. Hoag is an associate professor, and Herb A. Holloway is a former research assistant, Department of Agricultural and Resource, Economics North Carolina State University.

This research was funded in part by the North Carolina Agricultural Research Service.

The authors are grateful for comments and assistance provided by Douglas Young, Randy Rucker, and three anonymous *Journal* reviewers.

Dicks, Huang, Richardson et al.). However, the methodology used here reveals how farm characteristics such as program base and land productivity affect compliance incentives with detail not possible in a macrolevel analysis (Barbarika and Dicks, Huang) and with variety not possible in a case-farm or composite-farm analysis (Richardson et al., Wollenhaupt and Blase). Batie and Sappington examined the impacts of only a hypothetical cross-compliance program on Tennessee farms and did not identify the impacts of farm characteristics or distinguish between commodity program cross-compliance and conservation compliance.

The magnitude of CC's impact on crop production or soil erosion in the United States is not estimated here. Nevertheless, this study shows how certain farm characteristics may alter the effectiveness of CC in preventing soil erosion, thus indicating what it will accomplish. This study also shows how conservation compliance can fail to completely remove or out-compete continued incentives from commodity programs to use erosive farming systems.

Conservation Legislation

The conservation reserve program (CRP) provides annual rental payments for ten years on highly erodible land retired from production. Producers also receive 50% cost sharing to establish a vegetative cover. Acres entered into CRP reduce a farm's cropland base by the ratio of CRP to all cropland on the farm. Crop base is used to determine commodity program benefits.

The conservation compliance provision revokes eligibility for U.S. Department of Agriculture (USDA) farm program benefits to anyone producing an "agricultural commodity" on "highly erodible land" without an approved conservation plan. These benefits include price and income supports, disaster payments, crop insurance, Farmers Home Administration loans, Commodity Credit Corporation storage benefits, farm storage facility loans, and other programs under which payments are made for commodities produced by the farmer.

Highly erodible land is defined as having an erodibility index (EI) greater than or equal to eight ($EI = RKLS/T \geq 8$). The tolerance level T is the erosion level thought to maintain cropland productivity. The $RKLS$ factors measure the vulnerability of soil to erosion, as described in the universal soil loss equation, based on rain-

fall (R), erodibility (K), slope length (L), and slope gradient (S) (Wishmeier and Smith).

A producer with highly erodible land can obtain a conservation plan that will allow him to continue receiving program benefits by requesting one from the Soil Conservation Service (SCS). To facilitate the implementation of conservation compliance, the SCS developed standardized Alternative Conservation Systems (ACS). These systems are designed to meet conservation compliance requirements without extensive field visits (Ervin, Johnson and Clark). Landowners may select an ACS, which may differ by county and by soil capability subclass, or they can develop a customized plan with the assistance of SCS personnel.

The ACS establishes an allowable erosion guideline. A customized plan can be any combination of practices that reduces annual erosion rates to a comparable level provided by any of the ACS's. The customized plan must have a "CP" product less than or equal to the highest CP product of any of the pre-approved alternative conservation systems. The C (cover management) and P (supporting practice) factors measure the erosiveness of a particular system. When combined with the erodibility of the land, $RKLS$, the CP product indicates the erosion rate of a given system on a given soil.

Production Under Commodity and Conservation Programs

In the absence of government programs, net returns equal crop sales minus production costs. Commodity programs complicate production decisions through price and income supports and with cross-compliance cropping restrictions and temporary acreage set-aside requirements (Glaser). Under cross compliance, producers who receive program payments cannot produce program crops such as wheat or corn on more than their allotted base less an acreage set-aside. Base is the historical acres planted to a crop and set-aside is base acreage with no crop production for one year. Only nonprogram crops such as soybeans, hay, and sunflowers can be produced on more land than a producer's base. Therefore, the commodity programs effectively restrict crop production on both base and nonbase acreage.

A producer incurs at least two kinds of opportunity costs when using commodity programs. First, cropping restrictions may force production of less profitable crops on nonbase acres. Second, a temporary set-aside often is re-

quired in which no production is allowed. In return, the farmer effectively receives a guaranteed income for crops on his base acres equal to his farm or county's program yield times a target price (table 1).

The influence of program requirements on crop mix and other farm decisions can make production decisions complex (Perry et al.). However, conservation compliance and the conservation reserve make production decisions even more complicated. Conservation reserve adds rental payments, reduces base acreage and subjects a grower to a penalty if he fails to comply with conservation compliance over the ten-year period. To reduce erosion, CC adds new restrictions on crop type and production technology to those already imposed by cross compliance. Furthermore, conservation structures may be required and transactions costs are increased because a conservation plan must be filed. This added complexity increases the chance that conflicting production signals will be offered by cross compliance and CC.

The Model

A mixed integer programming model was used to determine optimal participation and cropping solutions for seventeen case study farms in Stanly County, North Carolina. A simplified version of the MIP tableau is given in table 2. The tableau represents a single farm unit as recorded by the Stanly County Agricultural Stabilization and Conservation Service (ASCS). Some farmers operate multiple farm units, but each is treated as a single entity by ASCS. Each farm has a

historical program acreage base in corn, wheat, or other crops.

Participants' farms have established program yields for each crop on which payments are made or they use a county average yield. The program yield is assumed equal to the average crop yield on the study farm for this analysis. Yield estimates by field were based on the distribution of soil mapping units by field and their relative productivity.

The model maximizes net returns to land, risk, and management for a single year by selecting the most profitable combination of cropping systems on each field on each farm while meeting appropriate constraints (table 2). A zero-one conservation compliance activity switches the CC program and commodity program constraints on and off. MIP can be used to solve endogenously for optimal cropping decisions when choices involve binary constraints. In this case, cropping decisions affect CC, which turns erosion and cross-compliance constraints as well as deficiency payment program benefits on or off.

The tableau in table 2 is divided into an objective row, blocks of constraint rows, and columns of input and output activities. Constraint blocks are defined for land use, commodity programs, soil erosion, equipment use, and returns from crop sales and CRP rent. The activities include conservation compliance (0 for compliance, 1 for noncompliance), specialized equipment use (integer), crop production, crop sales, CRP rents, and deficiency payments. Production is delineated by field because each has a unique productivity and susceptibility to soil erosion. For simplicity, only one piece of equipment, two crops, and two fields are depicted in table 2.

In the objective row (1), the costs of equipment use are added to annual operating costs of all crops on all fields and to the annual cost per acre of land enrolled in the Conservation Reserve Program (CRP). Returns include crop sales from all fields, CRP rental payments, and deficiency payments.

Production is constrained in row 2 such that all crops on all fields plus CRP acreage does not exceed available cropland. Rows 3 and 4 assure that the production on any field does not exceed its acreage. Rows 2, 3, and 4 will contain slack if the farmer is in conservation compliance because some land will be set aside. However, these constraints are binding when the producer does not comply or participate in commodity programs.

Table 1. Price and Program Assumptions for the Average and Forecasted Model Scenarios

	Static Scenario ^a	Forecasted Scenario
Corn market price	\$2.60/bu.	\$2.23/bu.
Corn target price	2.84	2.75
Corn set-aside	10%	10%
Wheat market price	3.57	3.26
Wheat target price	4.10	4.00
Wheat set-aside	10%	5%
Oat market price	2.20	1.83
Oat target price	1.49	1.44
Oat set-aside	10%	10%
Soybean market price	6.85	6.22

Static scenario sets all variables equal to the 1988 levels, and the forecasted scenario is based on FAPRI forecasts.

Table 2. Simplified Tableau of Mixed Integer Programming Model for Cross- and Conservation Compliance Decisions

Row	CC 0 = Comply 1 = Out	Equip. use (Integer)	Field 1			Field 2			Crop Sales			CRP		Deficiency Payment		RHS
			Crop 1 (ac)	Crop 2 (ac)	CRP (ac)	Crop 1 (ac)	Crop 2 (ac)	CRP (ac)	Crop 1 (\$/bu)	Crop 2 (\$/bu)	MP1	MP2	Rent (\$/ac)	Crop 1 (\$/ac)	Crop 2 (\$/ac)	
1 Objective		-OC	-PCI	-PC2	-C	-PCI	-PC2	-C					R	DP1	DP2	MAX
2 Land use			1	1	1	1	1	1	1							≤A
3 Land use			1	1	1	1	1	1								≤A1
4 Land use																≤A2
5 Commodity program	B												B/A	1 (1 - SA1)	1 (1 - SA2)	≤B
6 Commodity program	-M		1 (1 - SA1) -1	1 (1 - SA2)	1	1 (1 - SA1) -1	1 (1 - SA2)	1								≤A
7 Commodity program														1		≤0
8 Commodity program															1	≤0
9 Erosion (CC)	-M		CP	m	0											≤L
10 Erosion (CC)	-M					0	0	1								≤0
11 Equipment		-U	H1	H2		H1	H2									≤0
12 Crop sales			-Y1	-Y2		-Y1	-Y2		1	1						≤0
13 Crop sales																≤0
14 CRP rent					-1			-1					1			≤0

Note: This example includes only two crops and field combinations and only one type of equipment. Actual farms contained many fields, with several cropping options. Variables names:

- CC Conservation compliance
- OC Annual ownership cost of equipment
- PC Per acre operating cost, crop 1, crop 2
- MP Market price/unit, crop 1, crop 2
- DP Deficiency payment/acre, crop 1, crop 2
- M An arbitrarily large number
- m M/A
- SA Set aside rate for crop 1 or crop 2
- B Sum of all base acreage
- CP Product of "C" and "p" erosion factors per acre
- L "CP" limit under CC on field 1
- A Total crop acreage
- A1, A2 Acres in field 1 or 2
- U Maximum available hours of equipment use
- H Hours of equipment use on per acre of crop 1 or 2
- Y Yield per acre of crop 1 or 2
- C Conservation reserve program cost per acre
- R Conservation reserve program rent per acre

Commodity program and CRP constraints are defined in the commodity program block, rows 5–8. If a producer is participating in commodity programs and therefore in CC, column 1 is set to zero and row 5 requires that the deficiency payments are not paid on more acres than the base, less set-aside and CRP adjustments, allows. However, if a farmer does not participate, column 1 is set to 1 and constraint 5 would not allow the model to select any positive amount of crops with deficiency payments or CRP.

In row 5, CRP acreage reduces available base at a rate equal to $CRP / (\text{total cropland acres})$ as required by law. Base is reduced through the CRP rent column as the program enters CRP acres. For example, 100 acres of CRP on a 400-acre farm with 100 acres of corn base would reduce the base by 25% ($100/400$) to 75 acres. Program crops consume base at a rate of $1/(1 - SA)$, where SA is the annual set-aside requirement ranging from 0 to 1.0. Each acre of program crop would utilize 1.25 acres of base with a 20% set-aside ($1/(1 - .20)$); therefore, 60 acres of corn would require 75 acres of base. Program crops are constrained by a producer's current base and restrictions from CRP enrollment.

The crop and CRP base acreage reductions are contained in a single constraint. Base swapping is allowed among program crops to meet conservation compliance (ASCS). Therefore, only the total base, equal to all crop bases less CRP reductions, is constrained.

Row 6 assures that available cropland is not exceeded when a producer participates in the programs. If a producer does not participate, the compliance variable becomes 1, which when multiplied by the large negative number ($-M$), relaxes the constraint. Rows 7 and 8 require that an acre of production, with its associated costs, be brought into production for each acre of deficiency payment earned.

The erosion block is controlled by the conservation compliance integer variable. It is non-binding when a producer does not wish to receive deficiency payments or enroll land in CRP. When the constraint is binding, the CP factor of a crop system is limited to the maximum allowed by CC on the field where the crop is produced. Erosive systems that are not permitted on highly erodible fields are excluded by assigning them an arbitrarily large number, m , in the erosion block. For example, cropping system 2 is not allowed on field 1 and thus an m is entered in row 9. However, the same system is allowed on field 2 because the field is not highly erodible and erosion is not constrained.

Similarly, CRP is allowed on field 1 but not on field 2 because field 2 is not highly erodible.

In the intermediate run, equipment ownership costs are variable. Therefore, the cost of specialized equipment that changes between cropping systems was added into the tableau. Equipment was allowed to enter only in whole units. Therefore, each piece of equipment was more costly if it was not used to its full capacity. A unit of equipment relaxes the equipment constraint by the number of hours it can be used in a single year, U . Each acre of crop consumes these hours at a specific rate, H . When the constraint is binding, another whole unit of equipment is needed to allow more production.

In the final block, CRP acres are transferred into rent and yield is transferred from the field to sales activities. Yields were allowed to vary by field for each crop based on the soil types within the field.

A Case Study

A case study in Stanly County, North Carolina, is used to examine the impact of CC and commodity programs on net returns and soil erosion. The net returns to program participation with and without CC and under two levels of program benefits and price expectations were compared across seventeen actual farms. Net returns were optimized using mixed integer programming (MIP) for each farm.

This study is micro-oriented because returns are maximized on individual farms with a unique base acreage, soil productivity, number of fields, and size of operation. The most profitable combination of crops, production technology, and participation is solved field by field for each farm. However, the study also provides information about the broader impacts of CC by comparing the results of the seventeen farms to determine which characteristics influenced the participation decision, soil erosion, and net returns.

Data and Assumptions

The seventeen sample farms were selected from a randomly drawn data set for the Piedmont region of North Carolina. Stanly county is one of three counties drawn by lot. Each county had a probability of being selected equal to its relative share of corn production to all counties being considered in the region. Thirty-six farms were selected from those that had corn base recorded

in the Stanly county ASCS files. Samples were chosen randomly in four farm-size categories: 20–50 acres, 51–100 acres, 101–250 acres, and more than 250 acres. No effort was made to identify additional farms operated by the producer listed; thus, each farm number was treated as an independent entity. The final seventeen farms chosen reflected a variety of base, farm size, and soil productivity (table 4).

ASCS field maps were compared with soil survey maps made by the Soil Conservation Service (SCS) to determine the proportions of various soil types in each field. These estimates were combined with yield and erosion information to compute weighted average yields and erodibility for each field. Of 133 fields on the seventeen sample farms, 18% were highly erodible, 79% were potentially highly erodible. Only 3% were slightly erodible. SCS in North Carolina considers all potentially highly erodible soils highly erodible unless contested by the producer; then a field visit to measure slope is required. Therefore, the majority of fields in Stanly County are classified as highly erodible.

The Stanly County SCS provides farmers with a list of five alternative conservation systems (ACS) from which to choose (table 3). Farmers are free to choose from the list or to develop a customized system with a *CP* factor less than ACS's maximum *CP* factor. Stanly County does not offer any conservation structures such as terracing under any of their ACSs. However, a farmer could choose such a system if it reduces erosion by as much as the most erosive ACSs.

Enterprise budgets were developed for the ACSs and for commonly used systems that would not qualify under CC. Net profitability of a system equaled gross production returns less the variable production costs. Yields varied by field according to soil type. Production costs did not vary by field.

Farmers who participate in CC are required to establish and maintain permanent vegetative cover on all "critical areas" (determined by SCS).

Critical areas are draws or other areas where concentrated water flow increases the risk of soil erosion. They account for an estimated 6% of a typical field in North Carolina (B. Brock, North Carolina SCS, personal communication 1988).

Commodity prices and program provisions used in the MIP model were taken from the *FAPRI Review* (Food and Agriculture Policy Research Institute). Two scenarios were used: (a) a static scenario, where expected commodity prices, target prices, and set-aside requirements equal current levels; and (b) a forecasted scenario, where commodity prices, target prices, and set-aside requirements are equal to an average of the 1989/90 through 1994/95 FAPRI predicted prices (table 1).

Farmers are free to reallocate their existing acreage bases among most program crops (except cotton) on a one-time basis to meet the requirements of a conservation plan. For example, a corn base of 100 acres could be traded for 50 acres of wheat base and 50 acres of corn base. The 1988 bases for the seventeen North Carolina farms ranged from 32% to 100% (table 4).

Deficiency payments and CRP rental payments were the only government benefits considered. It is difficult to value other program benefits and these other benefits are little used in Stanly County. For example, in 1988, only one farmer used federal crop insurance; only five farmers bought crop insurance in 1989 (following a drought). The Farmers Home Administration issues only about five production loans annually.

Results

The MIP model optimizes returns for each farm under the following scenarios which have been designed to compare the impacts of cross and conservation compliance with various price and program scenarios: (a) static prices and programs without conservation compliance, (b) static

Table 3. Stanly County Alternative Conservation Systems

Number	Soils	Year 1	Year 2	CP-Factor ^a
1	GeB, GfB2, HeB	CC-CSG	NTSB	0.182
2	all others	NTC-CSG	NTSB	0.129
3	all others	NTC	NTSB	0.175
4	all others	NTSB-CSG	NTSB	0.140
5	all others	CC-NTSG	NTSB	0.162

^a The CP-factor was computed by the North Carolina State Office of SCS, and CC is conventional tillage corn/grain sorghum; CSG, conventional tillage small grain; NTSB, no-tillage soybeans; NTC, no-tillage corn/grain sorghum; NTSG, no-tillage small grain.

Table 4. Farm Base, Average Corn Yield, and Net Returns with and without Conservation Compliance Under Static and Forecasted Expectations

Sample Farm No.	Percent Base	Avg. Corn Yield (bu/ac)	Net Farm Returns to Management, Land, and Risk ^a			
			Static Expectations		Forecasted Expectations	
			Commodity	+ Cons. Comp	Commodity	+ Cons. Comp
1	32	86.2	3661	3661	2391	2391
2	37	90.0	7114	7114	5295	5295
3	46	83.5	41018	41018	27613 ^b	26735
4	54	69.1	6376	6376	4528 ^b	4407
5	56	87.1	28950	28950	21599 ^b	20648
6	59	73.1	11135	11135	7407 ^b	6781
7	71	72.5	1704	1704	1260 ^b	1114
8	72	87.8	24691	24691	20319 ^b	17410
9	73	89.2	6591	6591	5534 ^b	4847
10	80	83.5	1805	1805	1456 ^b	1202 ^b
11	81	80.3	5611	5611	4532 ^b	3915 ^b
12	82	66.9	2313	2313	1764 ^b	1667
13	91	75.6	2521 ^b	2494	2179 ^b	1780 ^b
14	94	77.0	5738 ^b	4987 ^b	4831 ^b	4070 ^b
15	100	86.7	20038 ^b	18567	18126 ^b	16402 ^b
16	100	76.9	1776 ^b	1513 ^b	1525 ^b	1248 ^b
17	100	76.0	7198 ^b	7044 ^b	6102 ^b	5240 ^b

^a Returns are also to equipment ownership that did not vary with system choices.

^b Participated in commodity programs.

prices and programs with conservation compliance, (c) forecast prices and programs without conservation compliance, (d) forecast prices and programs with conservation compliance.

Participation and Net Returns

Program participation and net returns for each farm under each scenario are shown in table 4 and table 5. The farms are ordered in table 4 on the basis of increasing proportions of base acreage to total crop acreage. Participation is more profitable as base increases under all four scenarios

but is relatively less profitable with CC. Farms 14, 16, and 17, which have a relatively large share of program base, profit from participation under all four scenarios, and farms 1 and 2, which have very little base acreage, are always better off not to participate. The profitability of participation on the other farms is affected by the scenario under which the model is run; that is, whether price and program benefit expectations are high or low.

Farms that have a high ratio of base to non-base acreage have relatively more benefits to protect than other farms but face similar costs to adopt conservation technology since all highly

Table 5. Effect of Conservation Compliance on Net Returns, Commodity Program Payments, and Erosion on All Seventeen Stanly County Farms

Scenario	Use Commodity Programs ^a (No. Farms)	Net Returns (\$)	Commodity Program Payments (\$)	Soil Erosion (CP-factor)
Static w/o cc	5	178,050	6,954	210
Static w/cc	3	175,668	2,896	207
% change	-40	-1.3	-58.4	-1.4
Forecasted w/o cc	15	136,288	45,154	381
Forecasted w/cc	7	125,010	31,050	135
% change	-53	-8.3	-31.2	-64.6%

^a Number among the seventeen farms participating in commodity program.

erodible fields on the farm must have a conservation plan. In addition, farms with a high average yield have more program benefits to protect. Therefore, farm losses stemming from CC, equal to net returns before CC minus net returns after CC, were regressed on the relative farm base (*FB*) and corn yield (*CY*). The results yielded a strong relationship as follows:

$$NR = -32.47 + 0.17(FB) + 0.32(CY),$$

(5.66) (4.00)

where *NR* is the difference in net returns, before and after CC, the terms in parentheses are *t*-values, *FB* is base/(crop acreage), and *CY* is the bushels of corn yield per acre; R^2 is 0.76. There is a positive relationship between both average yields and the proportion of base with the net farm losses stemming from CC. Producers possessing relatively small amounts of base can minimize losses from CC by discontinuing participation in commodity programs. However, those farmers with larger bases stand to lose more from CC either by sacrificing more program benefits if they do not comply or by adopting the necessary conservation measures. Farms already not participating did not realize any change in net returns (table 4). The overall effect on net returns for all seventeen farmers was an average reduction of 1.3% with static expectations and 8.3% with forecasted expectations (table 5).

Farmers' expectations about prices and program benefits influence program participation. Without conservation compliance, five farms (29%) maximized profits by participating in commodity programs under the static scenario, whereas fifteen farms (88%) maximized profits with program participation in the forecast scenario (table 5). The difference can be attributed to higher deficiency payments for all crops, a lower wheat set-aside rate, and a lower market price for oats under the forecasted scenario.

Under conservation compliance, participation in commodity programs is reduced in both scenarios, from 29% to 18% (5 to 3) and 88% to 41% (15 to 7) in the static and forecast scenarios, respectively (table 5). The lower commodity prices in the forecast scenario increase the benefits and lower the costs of conservation compliance. Therefore, program participation is higher under that scenario with or without conservation compliance.

Overall, deficiency payments to all seventeen farms were six to ten times higher under the forecast scenario compared to those under the static scenario (table 5). CC reduced deficiency

payments by 58% in the static scenario and 31% in the forecast scenario.

Fifteen of seventeen farms participated in commodity programs under the forecast scenario without CC, and four of these had some CRP acreage (Holloway). Conservation compliance caused eight of these farms to discontinue participation in the commodity programs. Six of these eight placed or increased acreage in CRP. The other two participated neither in CRP nor in commodity programs. Farms placing acreage in CRP must meet CC requirements on all of their fields, even though some no longer participate in commodity programs. For some of these farms, it was more profitable to take land out of production and put it into CRP than to plant it to a conservation crop. For others, it was more profitable to expand acreage beyond that allowed by their base, thus disqualifying them for commodity benefits (cross compliance) but allowing them to retain CRP benefits if crops produced on other fields were not erosive.

Soil Erosion

Soil erosion is affected indirectly on base and nonbase land through cross-compliance restrictions on cropping alternatives. For example, corn is a highly profitable crop that can be produced on base, and soybeans is the most profitable crop allowed on nonbase land. However, corn and soybeans can be highly erosive compared to wheat or wheat doubled cropped with soybeans. Therefore, the commodity programs can increase erosion on base land when corn is made more profitable than less erosive crops, and they can increase erosion on nonbase land by substituting erosive soybeans for small grains.

In Stanly County, the commodity programs increased soil erosion as described above. The forecast scenario provided greater program benefits than the static scenario and led to greater participation and soil erosion. As shown in table 5, without CC the total level of soil erosion on all seventeen farms was more than 80% higher under the forecast scenario than under the static scenario (381 compared to 210). The participation rate in commodity programs was about three times higher (15 compared to 5). However, because more farmers participated in commodity programs under the forecast scenario, CC was much more effective in reducing erosion (64.6% compared to 1.4%).

The potential of cross compliance to increase erosion is highly site specific. In a Missouri

study, for example, the most profitable crop, with or without programs, was found to be soybeans. Therefore, commodity programs would not necessarily encourage additional erosion on non-program lands because continuous soybeans would have been produced anyway (Wollenhaupt and Blase).

Concluding Comments

One of the factors influencing participation in future commodity programs is the cost of conservation compliance (CC) compared to the benefits of the commodity programs. Of the seventeen North Carolina farms reviewed, participation dropped by as much as half when conservation compliance was required. Actual sign-up rates for conservation plans have been high (Brock). These plans, however, require no immediate action and yet enable farmers to participate in the commodity program through 1995. Therefore, they may not accurately reflect actual intentions.

This analysis indicated that acreage base and soil productivity will have a substantial impact on program compliance. Farms with relatively little base acreage and relatively low yields likely will find it more costly to participate in commodity programs under CC than to quit using the programs altogether. Because all highly erodible fields need conservation plans, these farms have relatively less benefits to protect but face similar costs to comply.

The impacts of two program restrictions, cross compliance and CC, were examined. Cross compliance tended to increase soil erosion by restricting program crops to base acres. This forced farmers to produce more erodible and less profitable crops on nonbase acres. In Stanly County, continuously tilled soybeans were the most profitable crop allowed on non-base land and were more erosive and less profitable than soybeans double cropped with wheat. Double cropped wheat and soybeans requires base, however. Cross compliance increased erosion by over 80% when participation increased from about 30% of farms to almost 90% under the static and forecast scenarios, respectively. Conservation compliance more than offset the increase, however.

Across all farms examined, net farm income fell by 1% to 8% under CC depending on price and program expectations. Commodity program payments fell by 31% to 58%. Conservation compliance incentives to reduce soil erosion competed with the erosion-inducing effects of

cross compliance. Soil erosion declined by only 1% in the static price and program expectations scenario because commodity programs were not sufficiently attractive to induce producers to adopt needed conservation measures. A larger decline of 64% occurred in the forecast scenario because benefits were more attractive and therefore compliance remained stronger.

These results demonstrate how firm-level studies can provide insight that national studies may miss. Often the information gathered about the expected impact of a policy in aggregate analyses will not provide as much detail about individual farm impacts as models of the firm can reveal. Policy makers may find additional information similar to that provided here useful in dealing with individuals who are disproportionately affected by a given policy.

[Received August 1989; final revision received April 1990.]

References

- Barbarika, A., and M. Dicks. "Estimating the Costs of Conservation Compliance." *J. Agr. Econ. Res.* 40(1989): 12-20.
- Batie, S., and A. Sappington. "Cross Compliance as a Soil Conservation Strategy: A Case Study." *Amer. J. Agr. Econ.* 68(1986):880-85.
- Ervin, C. "Implementing the Conservation Title." *J. Soil and Water Conserv.* 44(1989):367-71.
- Food and Agriculture Policy Research Institute (FAPRI). *FAPRI Review—November 1988 10-year Baseline*. Review presented in Kansas City MO, Dec. 1988.
- Glaser, L. *Provisions of the 1985 Food Security Act*. Washington DC: U.S. Department of Agriculture, Econ. Res. Serv. Agr. Info. Bull. No. 498, 1986.
- Hertel, T., and P. Preckel. "Commodity-Specific Effects of the Conservation Reserve Program." *J. Agr. Econ. Res.* 40(1989):2-11.
- Holloway, H. *A Mixed Integer Programming Model for Analysis of the Impact of Conservation Compliance*. M.S. thesis, North Carolina State University, 1989.
- Huang, W. "Costs and Implications of Conservation Compliance." *J. Soil and Water Conserv.* 44(1989):521-26.
- Johnson, J., and R. Clark. "Conservation Compliance: What Loss with What Level of Farm Income Impact?" *J. Soil and Water Conserv.* 44(1989):458-61.
- Perry, G., B. McCarl, M. Rister, and J. Richardson. "Modeling Government Program Participation Decisions at the Farm Level." *Amer. J. Agr. Econ.* 71(1989):1011-20.
- Reichelderfer, K., and W. Boggess. "Government Decision Making and Program Performance: The Case of the Conservation Reserve Program." *Amer. J. Agr. Econ.* 70(1988):1-11.

- Richardson, J., D. Gerloff, B. Harris, and L. Dollar. "Economic Impacts of Conservation Compliance on a Representative Dawson County, Texas, Farm." *J. Soil and Water Conserv.* 44(1989):527-31.
- Wishmeier, W., and D. Smith. *Predicting Rainfall Erosion Losses from Cropland East of the Rocky Mountains*. Washington DC: U.S. Department of Agriculture, Agr. Res. Serv. Handbook No. 282, 1965.
- Wollenhaupt, N., and M. Blase. "The Economic Impact of Conservation Compliance: With Emphasis on Northern Missouri." Dep. Agr. Econ., University of Missouri, 1989.
- Young, E., and T. Osborn. *The Conservation Reserve Program: An Economic Assessment*. Washington DC: U.S. Department of Agriculture, Econ. Res. Serv. Agr. Econ. Rep. No. 626, 1990.

Measuring the Economic Efficiency of Producing Rural Road Services

Steven C. Deller and Carl H. Nelson

The research reported here examines the ability of a sample of Midwest township officials to produce low-volume rural road services in an economically efficient manner. Farrell-type measures of input use and scale efficiency are reported. Results suggest that over 50% of costs may be unnecessarily incurred because of input use inefficiency. Correlation between output measures and the efficiency measures suggests that larger jurisdictions are more efficient than smaller jurisdictions. In addition, 84.5% of the townships exhibit technology characterized by increasing returns to scale. These results suggest that jurisdictional consolidation of production-related responsibilities may yield substantial cost savings.

Key words: managerial efficiency, rural governments.

The efficient allocation of resources by local government remains an important issue, especially given the small scale of operation inherent to many rural governments. At issue is the ability of local public officials to control costs in the face of shrinking revenue sources. The research reported in this article provides some empirical evidence on the economic efficiency of a sample of small rural governments in the Midwest (Illinois, Minnesota, and Wisconsin) vested with rural road responsibilities.¹ Farrell-type measures of efficiency are computed for managerial and scale efficiency in order to test the ability of local public officials to provide rural road services in a least cost fashion.

The paper contains six sections. The first section describes managerial and scale efficiency within a public setting. The concepts of efficiency reference sets and efficiency measurement are then discussed. The network of low-

volume rural roads is then described, followed by the empirical results. The concluding section considers the study's policy implications.

Efficiency within a Public Setting

One view of local government maintains that the small-scale operation is inherently inefficient and costly (Chicoine and Walzer 1985, Bish and Osrom). Arguments are made that the physical consolidation of small government units would increase production efficiency, thus better utilizing available resources (Baumel and Schornhorst, Mercier). The consolidation argument traditionally hinges on scale economies in production. Indeed, empirical tests have generally indicated that scale economies are present in the production of public goods and services (Doeksen and Peterson, Fox).

A complementary argument addresses the ability of small, particularly rural, governments to develop an effective service provision program (Cigler; Reid 1984, 1986; Reid and Sullivan; Sokolow). The managerial capacity of smaller units of government to effectively perform functions is questioned. Managerial inefficiencies, if present, incur additional costs and exacerbate potential fiscal constraints. While the proponents of consolidation appeal to the notions of poor managerial skills and size economies, arguments about managerial efficiency

The authors are assistant professors, respectively, Department of Agricultural and Resource Economics, University of Maine, and Department of Agricultural Economics, University of Illinois.

Maine Agricultural Experiment Station External Publication No. 1438. Support for this research was provided in part by the Maine Agricultural Experiment Station, University of Maine, the Illinois Agricultural Experiment Station, University of Illinois, and the National Center for Supercomputing Applications.

The authors thank David L. Chicoine, Wes Seitz, Fred Giertz, and the referees for many useful comments. Any errors are the sole responsibility of the authors.

¹ The sample is limited to three states to preserve homogeneity in institutional arrangements and topography.

are limited to logical speculation (Sullivan and Reid).²

This study evaluates the efficiency of local governments by employing the tools of modern production economics and efficiency measurement. The Farrell method requires a normative criterion (e.g., cost minimization or profit maximization) to judge individual observations. Individual observations are then compared to these normative efficient reference sets. Measures of efficiency then can be correlated with relevant policy-related variables. Of particular interest in this study is the relationship between township size and efficiency.

The traditional consolidation literature has almost exclusively examined technical relationships between governmental organization and the supply of public goods. Because demand considerations are ignored, the traditional empirical work supporting governmental consolidation may not be sufficient. This research is concerned solely with the examination of efficient production of local public goods. The vital distinction between provision and production must be maintained when addressing policy considerations.

The Structure of Economic Efficiency

To compute the managerial and scale measures of efficiency, a technical definition of the efficient reference sets must be established. Scale efficiency refers to the location of the observed input combination relative to the scale efficient output level associated with constant returns to scale.³ Färe, Grosskopf, and Lovell (FGL) suggest that one measure of scale efficiency compares the distance between the boundary of the input requirement set for the given observations and the boundary of the input requirement set associated with constant returns to scale.

To calculate this measure FGL suggest the

construction of two functions which map inputs and outputs onto the real number line:

$$(1) \quad F(y, x) = \min_{\lambda} \{\lambda \in R_+ | \lambda x \in V(y)\}, \quad \text{and}$$

$$(2) \quad W(y, x) = \min_{\lambda} \{\lambda \in R_+ | \lambda x \in V^c(y)\}.$$

Here $y \in R_+^m$ represents a vector of outputs, $x \in R_+^n$ represents a vector of inputs, $V(y)$ represents the subset of all input vectors which yield at least y , and $V^c(y)$ is the input requirement set under constant returns to scale.⁴ The functions give the maximum amount by which an input vector can be radially decreased without altering the given output level. The difference between the two functions is the structure imposed on technology: the function $F(y, x)$ is constructed on an unrestricted technology whereas $W(y, x)$ is constructed on a technology exhibiting constant returns to scale.

The measure of scale efficiency is the distance between the boundary of $V(y)$ and $V^c(y)$, or

$$(3) \quad S(y, x) = W(y, x)/F(y, x).$$

If $S(y, x) = 1$, the unrestricted technology is identical to the constant-returns-to-scale technology, and the combination (y, x) is scale efficient. If $S(y, x) < 1$, then the technology yielding $F(y, x)$ has characteristics different from the technology yielding $W(y, x)$, and the combination is scale inefficient. Note that $S(y, x) \geq 1$ because $V^c(y)$ being a subset of $V(y)$ implies $W(y, x) \leq F(y, x)$.

To construct a measure of managerial efficiency, information is required about the goals of local public officials. Given an Inman view of local governments, a reasonable normative criterion is cost minimization.⁵

The cost-minimizing set of input vectors at input prices w , where prices are strictly positive, and outputs y is defined as

$$(4) \quad CM(y, w) = \{x \in V(y) | w'x = C(y, w)\},$$

² The presence of scale economies does not imply the presence of managerial inefficiencies, i.e., capacity shortfalls. The traditional test for scale economies implicitly assumes away the capacity issue because no estimate is made of the resources wasted or costs incurred because of poor management practices. If costs are increased as a result of managerial shortfalls, the traditional empirical tests for scale economies may be capturing capacity-induced costs and incorrectly attributing them to scale economies.

³ The use of output levels associated with constant returns to scale as a reference set is consistent with the traditional consolidation literature which searches for the minimum of the long-run average cost curve.

⁴ The input correspondence is assumed to satisfy the following axioms (Shepard): (a) impossibility of free production; (b) the input correspondence is bounded for any x ; (c) inputs are strongly disposable (monotonicity); (d') $V(y)$ is closed; and (e) strong disposability of outputs. See Färe, Grosskopf, and Lovell for measures of efficiency under alternative assumptions.

⁵ Inman's view of local governments may be modeled as a two-step process. The first step corresponds to provisionary decisions (demand or public choice), while the second step refers to production related decisions. Given that most local officials operate in the second step, a reasonable criterion may be to produce the demanded level of the public good at the lowest possible costs, i.e., cost minimization.

where the cost function $C(y, w)$ is defined as

$$(5) \quad C(y, w) = \min_x \{w'x | x \in V(y)\}.$$

If an input combination x yielding output level y is not in the cost-minimizing reference set, it is considered managerially inefficient.

Farrell suggests that a natural measure of managerial efficiency is the ratio of minimum and observed costs:

$$(6) \quad O(y, x, w) = C(y, w)/w'x.$$

If $O(y, x, w) = 1$, then the observed input combination is consistent with cost minimization and the observation is managerially efficient. If $O(y, x, w) < 1$, then the observed input combination is inconsistent with cost minimization.

Empirical Methods

Following Farrell, Shephard, Afriat, and FGL, a nonparametric representation of the relevant efficient reference sets is employed in this research. An alternative approach relies on statistical methods to estimate a parametric representation of the relevant set (Førsund, Lovell, and Schmidt). The nonparametric method is employed because of its handling of multiple outputs and its computational ease relative to the parametric method.

Mathematical programming methods construct a piece-wise linear reference technology from a sample of observations. Allow M to denote the matrix of observed outputs and N the matrix of observed inputs. Assuming that the axioms on technology hold, the reference technology can be expressed as

$$(7) \quad V(y) = \{x | zM \geq y, \quad zN \leq x, \\ z \in R_+^k, \quad y \in R_+^m\}$$

where the vector $z = (z_1, \dots, z_k)$ denotes the input utilization rate, or intensity level, k is the number of observations, and m is the number of outputs.

Afriat, Koopmans, and later Grosskopf have shown that alternative structures can be placed on technology by imposing various restrictions on the vector of intensity parameters. Specifically, by requiring z_i to be non-negative, as in (7), constant returns to scale is imposed. [Thus, for notational consistency, $V(y)$ in (7) should be expressed as $V^c(y)$]. By imposing $\sum_i z_i = 1$ and $z \in R_+^k$, increasing, decreasing, or constant returns are allowed; and imposing the restrictions

$\sum_i z_i \leq 1$ and $z \in R_+^k$ limits technology to non-increasing returns to scale.

The computation of the scale efficiency measure requires solution of (1) and (2), or the solution of the linear programming problems

$$(8) \quad F(y, x) = \min_{\lambda, z} \{\lambda | zM \geq y; \quad zN \leq \lambda x; \\ \sum_i z_i = 1; \quad z_i \geq 0\}, \text{ and}$$

$$(9) \quad W(y, x) = \min_{\lambda, z} \{\lambda | zM \geq y; \\ zN \leq \lambda x; z_i \geq 0\}$$

for each observation. The ratio of the solutions to the LPs (3) provides a scale efficiency measure for each observation.

Computation of managerial efficiency requires the solution to (4) and (5), or the solution to the linear programming problem

$$(10) \quad C(y, w) = \min_{x, z}^* \{w'x^* | zM \geq y; \\ zN \leq x^*; \quad \sum_i z_i = 1; \quad z_i \geq 0\}$$

for each observation. By construction, the solution to (10) is the minimum cost of producing output y at prices w , $C(y, w)$, and $O(y, x, w) = C(y, w)/w'x$.

The measure of managerial efficiency can be interpreted as a percentage cost reduction that could be achieved. This is because the cost of employing the observed input combination, x , given factor prices, w , is $w'x$, the cost of employing the cost-minimizing combination is $C(y, w)$, and by construction, $C(y, w) = \lambda w'x$. Thus, λ is the percentage of actual cost at which the observed output could have been produced. Conversely, one minus λ is the percentage reduction in costs that could be achieved by eliminating managerial inefficiency. A similar argument applies to scale efficiency.

Sources of managerial efficiency gains will be investigated by comparing radial and nonradial measures of technical efficiency. The radial technical efficiency measure is the solution to

$$\min_{\lambda, z} \lambda,$$

subject to

$$zM - y \geq 0$$

$$zN - \lambda x \leq 0$$

$$\sum_j z_j - 1 = 0$$

$$z_j \geq 0 \quad j = 1, \dots, k,$$

$$\lambda \in [0, 1].$$

The nonradial technical efficiency measure is the solution to

$$\min_{z, \lambda_1, \dots, \lambda_n} \sum_{i=1}^n \lambda_i / \sum_{i=1}^n \delta(x_i),$$

subject to

$$zM - y \geq 0$$

$$zN - \lambda \odot x \leq 0$$

$$\sum_j z_j - 1 = 0$$

$$z_j \geq 0 \quad j = 1, \dots, k$$

$$\lambda_i \in [0, 1] \quad i = 1, \dots, n,$$

where $\delta(x_i)$ is equal to one if $x_i > 0$, zero otherwise, and \odot denotes component-wise multiplication of the λ and x vectors.

The radial measure constructs a reference isoquant and shrinks the observed input level, x , along a ray from the origin, until it touches the isoquant. The nonradial measure constructs a reference isoquant and shrinks the observed input level x by reducing each component of x as much as possible while remaining on the isoquant. These measures are significantly different when the reference isoquant has segments that are parallel to one of the coordinate axes, as in the case of Leontief technologies. The radial measure simply shrinks observations until they lie on such an isoquant. The nonradial measure shrinks observations until they reach the corner of Leontief-type isoquants. Examining each computed λ_i yields insight into the sources of inefficiency.

Low-Volume Rural Roads

Low-volume rural roads are the responsibility of township and/or county governments. In several midwestern states, townships provide and maintain the basic road network, while counties are responsible for a network of overlapping higher service routes. In the United States, approximately 14,400 towns and townships are responsible, on average, for thirty-eight miles of low-volume roads (Chicoine, Walzer, and Deller). The typical county, however, is responsible for ten times the mileage. The small scale of operation at the township level provides the motivation for considering jurisdictional consolidation for low-volume rural roads.

Five studies have examined scale economies

in producing road services. Swanson studied the costs of providing low-volume township road services in Illinois. Lamb and Pine studied scale economies of rural road services in Kansas, while Leshner and Mapp examined costs of producing county road services in New York. More recently, Deller, Chicoine, and Walzer and Chicoine, Deller, and Walzer examined cost structures of low-volume rural road services in the Midwest. Each found some evidence of scale economies in the production process. While each study has specific limitations, none considered managerial efficiency.

Scale efficiency measures calculated under an incorrect assumption of managerial efficiency will be biased because the two types of efficiency are confounded. If input use is inefficient, then unbiased measurement of size efficiency requires that managerial inefficiency be removed from the data.

The data used for this study are from a mail survey of township officials in three midwestern states (Illinois, Minnesota, and Wisconsin) (Chicoine and Walzer 1984). The sample contains 295 rural jurisdictions from Illinois, 127 rural Wisconsin jurisdictions, and 24 rural Minnesota jurisdictions, for a total sample size of 446 independent observations.

A representative township in the three-state area is responsible for 49.2 miles of road; the largest township has 126 miles of road, and the smallest only four miles. The distribution of road surface types reflects the rural environment: 46.9% of total mileage is gravel, 39.8% is low bituminous, and 13.3% is high bituminous.⁶ The economic difference of surface types is evident from maintenance costs: a mile of gravel road requires approximately \$1,890 annually, while a mile of high bituminous requires \$4,889 annually (Walzer and Chicoine). To allow for the heterogeneous nature of low-volume rural roads, a multiple output model is specified where output is the miles of gravel, low, and high bituminous roads.

Following Deller, Chicoine, and Walzer and Chicoine, Deller, and Walzer, four inputs are used in the production of low-volume rural roads: labor, road graders, single-axle trucks, and surfacing material. The labor input is the number of full-time-equivalent employees. Graders and trucks are measured by number employed. The material input is composed of the amount of sur-

⁶ Bituminous roads are constructed with crushed rock and oil. High bituminous roads are a higher service level and are generally more durable than low bituminous roads.

facing material purchased by the township on an annual basis. All input levels are observed directly from the mail survey.

Prices for each input are available from several sources. The price of labor came from the mail survey and is measured as an average annual salary. Prices for the two capital items were based on a fixed proportion of replacement cost. Replacement costs for these items were determined first by depreciating new equipment prices to reflect the age of the equipment owned by the township. Then, proxies of capital prices were assumed to equal a fixed proportion of the depreciated values based on schedules of average annual equipment expenses. The price of a unit of surfacing material is based on engineering estimates of material requirements for resurfacing projects. To account for spatial differences, the capital and material prices were adjusted using regional cost-of-living indexes.⁷

Empirical Results

Calculation of the scale, cost, and technical efficiency measures described above required the solution of four sets of 446 linear programming problems: one set to calculate $F(y, x)$, one to calculate $W(y, x)$, one to calculate the nonradial measure $[NR(y, x)]$ and a final set to calculate $C(y, x)$. For spatial considerations, townships were grouped by mileage, and subsample means of the two measures of efficiency are reported for each mileage category.

Turning first to managerial efficiency, the typical township contained in this sample exhibits a low degree of managerial efficiency ($O(y, x, w) = .3488$) (table 1). Only eleven townships were identified as perfectly efficient in input use [i.e., $O(y, x, w) = 1.000$].⁸ Only minor variation occurred across the state-grouped subsamples, but significant variation occurred across size-grouped subsamples. The Farrell method

suggests that the typical township may significantly reduce operating costs if cost minimization is achieved.

Insight into the potential cost reductions is provided by the two measures of technical efficiency reported in table 2. Based on a computed F -statistic ($F = 3.95$), there appears to be a significant difference between the radial and nonradial measures of technical efficiency in the case of rural roads (table 2). In every case, this difference was caused by the nonradial measure shrinking labor usage more than the radial measure.

Because capital items are typically expected to be used in fixed proportions, it was surprising to find that the radial-nonradial difference is always caused by the labor variable. That is, the reference isoquant is parallel to the labor axis, and the nonradial measure always shrinks labor usage more than the radial measure.

The radial-nonradial difference is small at large mileages and large at small mileages. This suggests that smaller townships employ too much labor to provide road maintenance services, while larger townships do not. Differences in the quality of labor or tasks assigned to employees do not contradict this result because the averaging of townships over size categories should "average out" heterogeneity in quality and/or job assignments. Similarly, this result is not weakened by possible problems in the measurement of input prices because neither measure of technical efficiency is dependent on price data. Thus, the evidence is strong and robust that smaller townships may experience cost inefficiencies because of excessive labor, while larger townships (i.e., more than 75 miles of road) do not employ excess labor, on average.

Turning to scale efficiency, the typical township contained in this sample was identified by the Farrell method as scale inefficient [$S(y, x) = .6912$] (table 3). Based on these computations, 53 townships (11.9% of the total sample) were scale efficient and 16 townships (3.6% of the sample) exhibited decreasing returns to scale. The remaining 377 townships (84.5% of the sample) exhibited increasing returns to scale, and smaller townships were more scale inefficient than larger townships. These measures of scale efficiency suggest that jurisdictional boundaries that limit the scale of production may increase operation costs unnecessarily.

While the evidence provided by the efficiency measures suggests that larger road jurisdictions may be more cost effective, the statistical significance of the difference in efficiencies has not

⁷ For more detail on input prices see Deller, Chicoine, and Walzer; Chicoine, Deller, and Walzer.

⁸ Following Varian, a simple nonparametric pretest of the data identified eleven townships as cost rational. Five of the eleven townships identified by the Varian pretest were identified as input use efficient using the Farrell approach.

These observations may constitute outliers in the data. To test the sensitivity of the computed measure of overall efficiency, these eleven observations were removed from the sample. The recomputed measure of overall efficiency for the typical township increased to .4745. A simple test of sample means suggests that this increase may be significant ($F = 22.38$). However, the basic policy conclusion remained unaltered.

Table 1. Nonparametric Measure of Managerial Efficiency

Town Grouping by Mileage Range	Illinois	Minnesota	Wisconsin	Three States
0-25	.2648	.3851	.2335	.2732
σ	(.14) ^a	(.43)	(.08)	(.18)
n	20	4	7	31
26-50	.2680	.5329	.3264	.2997
	(.16)	(.27)	(.19)	(.19)
	160	14	68	242
51-75	.3707	.3307	.3568	.3654
	(.17)	(.18)	(.18)	(.18)
	107	6	42	155
76-100	.6375		.5896	.6215
	(.26)		(.19)	(.24)
	18		9	27
100+	.7031		1.0000	.9152
	(.42)		(.00)	(.22)
	2		5	7
Total	.3267	.4577	.3795	.3488
	(.20)	(.29)	(.23)	(.22)
	295	24	127	446

^a Number in parentheses is the standard deviation of the mean followed by the number of townships in each category.

Table 2. Measures of Technical Efficiency

	Town Grouping by Mileage					
	0-25	26-50	51-75	76-100	100+	Total
$F(y, x)(\text{radial})$	1.0000	.9738	.9687	.9785	1.0000	.9749
σ	(.00) ^a	(.11)	(.12)	(.07)	(.00)	(.11)
n	31	242	155	27	7	446
$NR(y, x)(\text{nonradial})$.3438	.4183	.5578	.7780	.8579	.4857
σ	(.06)	(.15)	(.18)	(.20)	(.19)	(.20)
n	31	242	155	27	7	446

^a Number in parentheses is the standard deviation of the mean followed by the number of town ships in each category.

Table 3. Nonparametric Measure of Scale Efficiency

Town Grouping by Mileage Range	Illinois	Minnesota	Wisconsin	Three States
0-25	.4243	.4526	.4442	.4325
σ	(.23) ^a	(.27)	(.26)	(.24)
n	20	4	7	31
26-50	.6254	.7293	.6478	.6377
	(.17)	(.24)	(.19)	(.18)
	160	14	68	242
51-75	.7912	.7848	.7654	.7840
	(.13)	(.17)	(.16)	(.14)
	107	6	42	155
76-100	.9246		.9010	.9167
	(.12)		(.11)	(.11)
	18		9	27
100+	.9352		.8296	.8598
	(.09)		(.24)	(.21)
	2		5	7
Total	.6860	.6971	.7021	.6912
	(.20)	(.25)	(.21)	(.20)
	295	24	127	446

^a Number in parentheses is the standard deviation of the mean followed by the number of townships in each category.

Table 4. Statistical Tests of Equality across Mileage Ranges

	Anova F (Prob > F)	Wilcoxon χ^2 (Prob > χ^2)	Median χ^2 (Prob > χ^2)	van der Waerden χ^2 (Prob > χ^2)	Savage χ^2 (Prob > χ^2)
$O(y, x, w)$	35.07	74.43	45.72	84.74	116.49
(Managerial)	(.0001) ^a	(.0001)	(.0001)	(.0001)	(.0001)
$S(y, x)$	51.60	136.44	109.75	142.37	92.60
(Scale)	(.0001)	(.0001)	(.0001)	(.0001)	(.0001)
$F(y, x)$	0.65	3.79	0.00	3.20	3.92
(Radial)	(.6237)	(.4350)	(1.000)	(.5248)	(.4163)
$NR(y, x)$	3.95	29.45	21.55	38.31	14.32
(Nonradial)	(.0007)	(.0001)	(.0015)	(.0001)	(.0263)

^a Number in parentheses is the marginal significance level.

been established. A test of subsample mean equivalence (F -test), a nonparametric test of central tendency (median test), and several nonparametric tests which compare the distribution (Wilcoxon, van der Waerden, and Savage tests) of the measures across the subsamples were conducted. The results of these tests, reported in table 4, indicate significant differences across all mileages, except for the radial measure of technical efficiency. The strong positive relationship between size and overall efficiency supports the reform traditionalist hypothesis that larger townships are more efficient.

The impact of combining two average size townships can be examined using simple regression analysis. Based on the regression equations reported in table 5, a typical township with 23.1 miles of gravel roads, 19.6 miles of low bituminous roads, and 6.5 miles of high bituminous roads, has an expected overall managerial efficiency of .3608. If the production responsibilities

of two such townships were combined, the expected level of overall input use efficiency is .6184, an increase of nearly 42%. The positive statistical relationship between road mileage and managerial and scale efficiency further supports the increased efficiency of larger townships in producing road services.

Concluding Comments

The empirical evidence from a sample of mid-western townships suggests that larger townships are more efficient than smaller townships. In particular, smaller townships appear to employ an excess amount of labor. Efficiency measurement shows that over 50% of costs may be unnecessarily incurred by the typical smaller township because of managerial inefficiency. In addition, only 11.9% of the sample of townships were identified as scale efficient (i.e., exhibiting constant returns to scale), and 84.5% of the sample exhibited increasing returns to scale.

The correlation between efficiency and township size is strongly positive, and a simple regression of output levels on the measures of efficiency suggests that increasing jurisdictional size increases efficiency. These results suggest that substantial cost reductions could be realized by restructuring rural road maintenance.

These results should not necessarily be used to promote the physical consolidation of smaller jurisdictions into larger, more economically efficient units because they address only production efficiency. Provisionary impacts of consolidation have not been examined. Therefore, a responsible policy prescription supported by these results is consolidation of production responsibilities, something considerably short of jurisdictional consolidation.

Table 5. Regression Estimates of Mileage onto the Nonparametric Measures of Efficiency

	Dependent Variable	
	$O(y, x, w)$ (Managerial)	$S(y, x)$ (Scale)
Gravel miles	.00513 (.001) ^a	.00539 (.0005)
Low bituminous miles	.00372 (.001)	.00641 (.001)
High bituminous miles	.00846 (.001)	.00743 (.001)
Constant	.10217 (.026)	.39240 (.023)
F	48.22	66.05
R^2	.2466	.3095

^a Number in parentheses is the standard error of the estimate.

One means of production consolidation is contractual agreements between jurisdictions. Smaller units would preserve provisionary authority in arranging the terms of any agreement, but production responsibilities would belong to a larger, more efficient unit. Alternatively, smaller units may cooperate in contracting a private firm to produce the public good. Or, an engineer may be jointly hired by a union of local units to perform production related activities using public resources. Arrangements for more efficient allocation of rural road maintenance labor should be investigated because current allocation is highly inefficient.

[Received June 1989; final revision received March 1990.]

References

- Afriat, S. N. "Efficiency Estimation of Production Functions." *Int. Econ. Rev.* 13(1972):568-98.
- Baumel, P. C., and E. Schornhorst. "Local Rural Roads and Bridges: Current and Future Problems and Alternatives." *Transport. Res. Rec.* 898(1983):374-78.
- Bish, R. L., and V. Ostrom. *Understanding Urban Government*. Washington DC: American Enterprise Institute, 1972.
- Chicoine, D. L., S. C. Deller, and N. Walzer. "The Size Efficiency of Rural Government: The Case of Low-Volume Rural Roads." *Publius* 19(1989):127-38.
- Chicoine, D. L., and N. Walzer. *Financing Rural Roads and Bridges: In the Midwest*. Washington DC: Office of Transportation, 1984.
- . *Governmental Structure and Local Public Finance*. Boston: Oelgeschlager, Gunn, and Hain, 1985.
- Chicoine, D. L., N. Walzer, and S. C. Deller. "Financing Rural Roads and Bridges: Issues and Trends." *Profitability and Mobility in Rural America*, ed. W. R. Gillis, pp. 43-72. State College: Pennsylvania State University Press, 1989.
- Cigler, B. A. *Setting Smalltown Research Priorities: The Service Delivery Dimension*. Washington DC: U.S. Department of Agriculture, Econ. Res. Serv., 1987.
- Deller, S. C., D. L. Chicoine, and N. Walzer. "Economies of Scale and Scope in Rural Low-Volume Roads." *Rev. Econ. and Statist.* 70(1988):459-65.
- Doeksen, G. A., and J. Peterson. *Critical Issues in the Delivery of Local Government Services in Rural America*. Washington DC: U.S. Department of Agriculture, Econ. Res. Serv., 1987.
- Färe, R., S. Grosskopf, and C. A. K. Lovell. *The Measurement of Efficiency of Production*. Boston MA: Kluwer-Nijhoff Publishing, 1985.
- Farrell, M. J. "The Measurement of Productive Efficiency." *J. Royal Statist. Soc., Series A General* 120(1957):253-81.
- Førsund, F. R., C. A. K. Lovell, and P. Schmidt. "A Survey of Frontier Production Functions and of Their Relationship to Efficiency Measures." *J. Econometrics* 13(1980):5-25.
- Fox, W. F. "Size Economies in Local Government Services: A Review." Washington DC: U.S. Department of Agriculture, 1980.
- Grosskopf, S. "The Role of the Reference Technology in Measuring Productive Efficiency." *Econ. J.* 96(1986):499-513.
- Inman, R. R. "The Fiscal Performance of Local Governments: An Interspersive Review." *Current Issues in Urban Economics*, ed. P. Mieskowski and M. Strassheim. Baltimore MD: Johns Hopkins University Press, 1979.
- Koopmans, T. C. "Examples of Production Relationships Based on Microdata." *The Microeconomic Foundations of Macroeconomics*, ed. G. C. Harcourt. Boulder CO: Westview Press, 1977.
- Lamb, S. W., and W. H. Pine. "Influences of Size and Administrative Organization on Costs of Rural Roads." *S. J. Agr. Econ.* 6(1974):157-160.
- Leshner, W. G., and H. P. Mapp. "Economies of Size in Highway Maintenance and Administration: A Preliminary Analysis for the Counties of New York State." *Northeast. J. Agr. Econ.* 3(1974):90-111.
- Mercier, C. R. "Low-Volume Roads: Closure and Alternative Uses." *Transport. Res. Rec.* 898(1983):110-15.
- Oakerson, R. J. "Local Public Economics: Provision, Production and Governance." *Intergovernmental Perspectives* 13(1987):20-25.
- Reid, J. N. "Building Capacity in Rural Places: Local Views on Needs." *Perspectives on Management Capacity Building*, ed. B. W. Honadle and A. M. Howitt. New York: State University of New York Press, 1986.
- . *Rural Government Capacity: Institutional Authority and Local Leadership*. Washington DC: U.S. Department of Agriculture, Econ. Res. Serv., 1984.
- Reid, J. N., and P. J. Sullivan. "The Rural Infrastructure Problem: A National Perspective." *Financing Local Infrastructure in Nonmetropolitan Areas*, ed. D. L. Chicoine and N. Walzer. New York: Praeger Publishers, 1986.
- Shephard, R. W. *The Theory of Cost and Production Functions*. Princeton NJ: Princeton University Press, 1970.
- Sokolow, A. D. "Local Governments: Capacity and Will." *Nonmetropolitan America in Transition*, ed. A. H. Hawley and S. M. Mazic. Raleigh: University of North Carolina Press, 1981.
- Sullivan, P. J., and J. N. Reid. "Rural Infrastructure: A Look at Policy Issues." *New Dimensions in Rural Policy: Building Upon Our Heritage*. Washington DC: U.S. Congress, Joint Economic Committee, Subcommittee on Agriculture and Transportation, 1986.
- Swanson, E. R. "Rural Road Costs and Size of Road Unit." *Current Econ. Comment* 18(1956):25-32.
- Varian, H. R. "The Nonparametric Approach to Production Analysis." *Econometrica* 52(1984):579-97.
- Walzer, N., and D. L. Chicoine. *County Roads and Bridges: Finance and Administration*. Washington DC: U.S. Department of Agriculture, Office of Transportation, 1987.

Dietary Effects of the Food Stamp Program

Barbara Devaney and Robert Moffitt

Based on data from the 1979–80 Survey of Food Consumption in Low-Income Households, this paper estimates the effects of changes in cash income and the food stamp benefit on household nutrient availability, while controlling for two potential sources of selection bias. The major finding of the empirical analysis is that the estimated dietary effects of changes in food stamp benefits are considerably larger than those resulting from changes in cash income, with estimates of the ratios of the MPC for the food stamp benefit to the cash-income MPC ranging from three to seven across nutrients. No significant evidence of selection bias was found, and the estimated dietary effects of food stamp benefits from the selection bias models are similar to those from the basic model estimated by ordinary least squares regression.

Key words: food stamp benefits, marginal propensity to consume, nutrient availability, selection bias.

The objective of this paper is to assess the dietary effects of the Food Stamp Program (FSP). Despite a substantial literature on the food expenditure and dietary effects of the FSP, empirical results reported in the various studies differ. The study addresses the extent to which the FSP raises the quality of diets of participating households.¹ A related question is the marginal effect of food stamp benefits on diet quality. That is, what are the dietary effects of an additional dollar of benefits, and do coupon benefits affect diet quality more than equivalent cash income?

Model

It is assumed the household chooses from J foods Q_j , $j = 1, \dots, J$, and a composite nonfood good C . The utility function is

$$(1) \quad U(Q_1, Q_2, \dots, Q_J, C).$$

Barbara Devaney is a senior economist, Mathematica Policy Research, and Robert Moffitt is a professor, Department of Economics, Brown University.

This research was funded by the Food and Nutrition Service of the U.S. Department of Agriculture and contract FNS-53-3198-5-47. Any opinions expressed herein are those of the authors and not of the Food and Nutrition Service, Mathematica Policy Research, or Brown University.

The authors gratefully acknowledge the thoughtful comments of Steven Cole, Tom Fraker, and three anonymous reviewers. Any remaining errors are the responsibility of the authors.

¹ The discussion here presumes that one overall measure of household diet quality is available, although a variety of dietary outcome measures have been used in previous analyses of the FSP.

Let Y be total cash income, excluding food stamp benefits; B , the food stamp benefit; and P_j , the price of food good j relative to the price of C . The budget constraint can be written as

$$(2) \quad Y + B = \sum_{j=1}^J P_j Q_j + C.$$

In other words, income plus food stamp benefits is spent on the J food goods and nonfood consumption.

Maximizing utility subject to the constraint on household resources leads to J different demand functions for the food goods which can be written in the form:

$$(3) \quad Q_j = f_j(P_1, P_2, \dots, P_J, Y, B), \quad j = 1, \dots, J.$$

Cash income (Y) and food stamp benefits (B) are entered separately into the demand functions to allow for different effects of changes in cash income and the food stamp benefit levels on food consumption.

To include nutrients in this model, assume that there are K nutrients N_k , $k = 1, \dots, K$, and that each unit of food good yields a_{kj} of nutrient N_k . The K nutrient equations can be written as follows:

$$(4) \quad N_k = \sum_{j=1}^J a_{kj} Q_j, \quad k = 1, \dots, K.$$

Equations (3) and (4) constitute the "true" model of the determinants of nutrient levels. An increase in income or the food stamp benefit, for example, increases the quantity of each food good consumed (Q_j), though some may fall if they are inferior goods. An increase in a food good raises the availability of each nutrient in household diets, the amount depending upon the magnitudes of the a_{kj} 's.

Econometric Estimation

The usual approach to estimating the model depicted by equations (3) and (4) is to focus on the food good demand functions, (3). Most commonly, these individual food-good equations are aggregated to a total food expenditure equation by multiplying the Q_j in each food-good equation by its price and adding them up across goods. However, an equation for total food expenditures does not allow the use of equation (4) to determine the nutrients from food because there is no exact relationship between an increase in total food expenditures and an increase in each of the nutrients; the relationship depends upon the combination of the individual foods embodied in the rise in food expenditures.

Another approach to estimating the theoretical model has been to estimate relationships between nutrient levels and household food expenditures. This is essentially a recursive model in which cash income and the food stamp benefit are assumed to affect household food expenditures, and household food expenditures influence nutrient availability levels. The effects of the FSP, for example, are then traced through their impacts on food expenditures which, in turn, affect nutrient availability levels. However, this approach leads to model specification bias since there is no fixed relationship between household food expenditures and each nutrient; that is, it is generally not possible to use equations (3) and (4) to derive equations with nutrient availability on the left-hand side and total food expenditures on the right. Specifically, the coefficient on total food expenditures in a nutrient equation would represent only some undetermined combined effect of income and food stamps benefits on nutrient levels.

This specification bias could have serious consequences for the estimation of the effect of food stamps on nutrient levels. When the food stamp benefit is entered separately into food expenditure equations, the results of past studies suggest that changes in FSP benefit levels have

a stronger effect on food expenditures than do changes in the level of money income (Chen, Smallwood and Blaylock). It is, therefore, likely that increases in food stamp benefits also have different effects on the consumption of individual food items than do increases in money income. In this case, the food expenditure coefficients in the nutrient equations will represent biased estimates of the effect of the food stamp benefit on nutrient levels, because those coefficients will only represent average effects of increases in money income and in the food stamp benefit.

Adding variables for income, for the food stamp benefit or FSP participation, or both, to the nutrient equations does not reduce the specification bias. In fact, the degree of bias probably increases because there is no reason for food expenditures to remain in the equation after income and FSP benefit variables are added directly.²

Proposed Estimation Procedure

Given the problems with estimating the nutrient equations with aggregate food expenditures in place of the individual food items, the major issue is how to estimate the dietary effects of the FSP. Because there are hundreds of individual food goods and thousands of a_{kj} 's the data do not allow the direct estimation of the individual food consumption equations. However, substituting the individual food demand equations (3) into the nutrient equations (4) results in a set of reduced-form nutrient equations of the form,

$$(5) \quad N_k = g_k(P_1, P_2, \dots, P_j, Y, B).$$

In the reduced-form equations, nutrient levels are related to prices, the food stamp benefit level, cash income, and other variables affecting the demand for food goods. Recognizing that price levels are generally constant for cross-sectional data, a linear regression equation of the form

$$(6) \quad N_{ki} = \alpha_k + \beta_k Y_i + \delta_k B_i + X_i \phi_k + \epsilon_{ki}$$

can be estimated for each nutrient k , where " i " denotes household i , Y_i is household cash income, B_i is the food stamp benefit, X_i is a set of other variables thought to affect the demand

² In addition to these difficulties, entering food expenditures on the right-hand side of the nutrient equations in (4) would lead to a type of simultaneous-equations bias, as food expenditures are endogenous. Therefore, the coefficient on food expenditures in a nutrient equation will be biased upwards.

for food goods and, hence, nutrients, and ϵ_{ki} is an error term. Coefficients β and δ represent the combined effects of (a) the effect of income and the food stamp benefit on the demand for food goods and (b) the effect of food consumption on nutrient levels. The two effects cannot be separated out, but this is unnecessary when determining the effects of the FSP. The important point is that the coefficients correctly capture the effects of income and food stamp benefits on nutrient availability working through the hundreds of individual food goods, even though those individual food consumption levels are not used in the estimation.

Equation (6) is straightforward and simple to estimate using ordinary least squares (OLS) regression. It is based on a framework in which food consumption decisions depend on income, prices, and other factors, and nutrient availability is a byproduct of the food consumption decisions.

Data and Empirical Results

The data used in this report are from the 1979–80 Survey of Food Consumption in Low-Income Households (SFC-LI). This survey was conducted from November 1979 through March 1980 using a national probability sample of approximately 2,900 low-income housekeeping households eligible to receive benefits under the FSP.³ Detailed information on household food use was collected by the SFC-LI. Household food use refers to food and beverages used from household food supplies during the seven days preceding the interview. Food purchased with cash, credit, or food stamps and food that was home-produced, received as a gift or payment for work, or received through other food programs are all included in the measure of household food use. In addition to the data on food use, information was obtained on household characteristics related to food use, such as participation in the FSP, participation in other food assistance programs (School Lunch, School Breakfast, and WIC), household consumption, income, education and employment of the household heads, urbanization, and tenancy.

Data on household food energy and nutrient availability are calculated from the quantity of each food item used from household food sup-

plies. Caloric and nutrient contents of each food item are obtained from tables of the nutritive value of foods. Total household availability of food energy and nutrients is derived by summing the food energy and nutritive values of the individual food items used. Nutritive values pertain to the edible portion of the food used from household food supplies, with some adjustments for vitamin losses during preparation.

Household nutrient availability data from the SFC-LI are based on food used from household food supplies. Nutritive values are not available for food eaten away from home. If the number of meals eaten away from home differs among groups of households, differences in nutrient availability will be observed regardless of whether or not any differences exist in the nutritive value of food used at home. Therefore, an adjustment for the proportion of meals eaten at home is needed when comparing nutrient availability from food used at home by subgroups of low-income households.

Table 1 shows the means of the major variables used in the empirical analysis. Eleven major nutrients were examined, and each is scaled by the number of equivalent nutrition units in the household. The number of equivalent nutrition units (ENUs) is one measure of household size and is defined as the number of adult equivalent males eating meals from household food supplies. It adjusts actual household size for both the age-sex composition of family members and guests and the proportion of meals eaten at home during the designated week. The adjustment procedure weights each household by (a) the nutritional requirements of the member relative to the nutritional requirements of an adult male aged 23–50, where the nutritional requirements are based on the 1980 Recommended Dietary Allowances (RDA) for each nutrient, and (b) the proportion of meals eaten at home during the week. This second part of the weighting scheme is necessary for analyses of nutrient availability since, as noted above, such nutrient data are based only on food used at home. Thus, the ENU adjustment is required not only for the differing age-sex compositions of each household but also because only food used at home is measured.

The income variables shown in table 1 are scaled by the number of adult male equivalents (AME), based on the 1980 RDA for food energy.⁴ Cash income is about eight times larger

³ Housekeeping households contain at least one person having 10 or more meals from household food supplies during the 7 days preceding the interview.

⁴ AME, rather than ENU, is used to scale the independent variables in the analysis because household size in equivalent nutrition units may be an endogenous variable since it depends on the proportion of meals eaten at home.

Table 1. Means of the Variables Used in the Analysis (N = 2,925)

Variable	Mean Value
<u>Nutrients per ENU per day^a</u>	
Food energy (Kcal)	3,988
Protein (g)	129
Vitamin A (IU)	11,414
Vitamin C (mg)	139
Thiamin (mg)	2.71
Riboflavin (mg)	3.23
Vitamin B ₆ (mg)	2.56
Calcium (mg)	1,000
Phosphorous (mg)	1,710
Magnesium (mg)	464
Iron (mg)	16.9
<u>Income per AME^b (\$ per week)</u>	
Cash income	\$47.23
Food stamp benefit ^c	5.42
Food stamp benefit-participants only	10.84
Subsidy value of school lunches	1.25
Subsidy value of school breakfasts	.17
Value of home-grown food	.53
Value of gift/pay food	.88
<u>Household characteristics</u>	
FSP participation rate	.50
AME (food energy)	2.63
Female head	.94
Guest meal per AME	.75
North central	.14
South	.67
West	.08
Spanish	.07
Suburban	.15
Nonmetropolitan	.37
Head of Household 35-59	.35
Head of Household 60+	.32
Black	.49

Source: 1979-80 Survey of Food Consumption in Low-Income Households, USDA.

^a ENU is equivalent nutrition units, which is the number of equivalent adult males eating from household food supplies.

^b AME is number of equivalent adult males in the household.

^c Includes zeros for nonparticipants.

than the average FSP benefit overall. However, the average weekly food stamp benefit per AME for participants only is \$10.84, which is roughly 28% of cash income for participating households (not shown).

The other variables shown in table 1 indicate that, on average, \$2.83 per adult male equivalent per week was from foods received either through the school nutrition programs, as gift or pay, or from home-grown food. The low-income sample was divided about evenly between FSP participants and nonparticipants, between blacks and non-blacks, and by age of the household head (<35, 35-59, 60+). In addition, average household size was 2.63 adult male equivalent

adults, the vast majority of the households had a female head present (94%),⁵ and the sample was located largely in the South and in rural or nonmetropolitan areas.

Table 2 shows the OLS estimates for each nutrient equation. These findings show that the household availability of all nutrients is positively associated with both the food stamp benefit level and money income. The most striking result is that the estimated marginal impacts of the food stamp benefit consistently and significantly exceed those of cash income. While the estimates indicate positive and statistically significant effects on nutrient availability for both the food stamp benefit and cash income, the coefficient on the money income variable is always less than the coefficient on the food stamp benefit, indicating smaller marginal effects of money income on nutrient availability.

This point is examined in more detail in table 3. The first two columns show the "MPCs" (i.e., the marginal propensity to "consume" nutrients for the food stamp benefit and for cash income). The cash income MPCs are evaluated at the mean of cash income, and therefore represent the average MPC in the sample. As the table indicates, the MPCs for the food stamp benefit are much greater than the cash income MPCs. The ratio of the MPC for the food stamp benefit to the cash income MPC is never less than 3 and is as high as 7. This large difference is discussed in more detail below.

To judge the relative size of the estimated dietary effects, the third and fourth columns of table 3 show the marginal effects of the food stamp benefit and cash income as a percentage of the adult male RDA. The percentage effects of changes in weekly cash income are quite low, ranging from .3% to 1.2% of the adult male RDA, while the percentage effects of changes in the weekly food stamp benefit are higher, ranging from 1.8% to 3.9% of the adult male RDA. Interestingly, the percentage effects of changes in the food stamp benefit are similar in magnitude for most nutrients. That is, nutrient availability increases from between 1.8% and 3.9% of the RDA for a one-dollar increase in the food stamp benefit. These findings imply that increases in the weekly food stamp benefit are generally allocated proportionally among the nutrients examined.

The final two columns in table 3 show the estimated total availability of nutrients (as opposed to marginal changes) attributable to the

⁵ "Female head present" refers either to households with a female head only or to households with both a male and female head.

Table 2. OLS Estimates of Equations for the Availability of Nutrients in Food Used from Home Food Supplies: U.S. Low-Income Households, 1979-80 (standard errors in parentheses, N = 2,925)

Explanatory Variables	Food Energy	Protein	Vitamin A	Vitamin C	Thiamin	Riboflavin	Vitamin B ₆	Calcium	Phosphorus	Magnesium	Iron
Constant	2,951 ^{a,b} (238)	104.79 ^{a,b} (7.17)	7,640 ^{a,b} (1,230)	110.85 ^a (13.45)	1.890 ^{a,b} (.166)	2.579 ^{a,b} (.194)	2.355 ^{a,b} (.155)	907 ^{a,b} (73)	1,518 ^{a,b} (104)	407.7 ^{a,b} (27.9)	13,168 ^{a,b} (1,093)
Household weekly food stamp benefit, per adult male equivalent	52 ^{a,b} (6)	1.81 ^{a,b} (.18)	156 ^{a,b} (31)	1.97 ^{a,b} (.41)	.040 ^{a,b} (.004)	.052 ^{a,b} (.005)	.039 ^{a,b} (.004)	18 ^{a,b} (2)	30 ^{a,b} (3)	7.1 ^{a,b} (.7)	.387 ^{a,b} (.044)
Household weekly money income per adult male equivalent ^a	16 ^{a,b} (4)	.59 ^{a,b} (.11)	39 ^{a,b} (20)	.90 ^{a,b} (.25)	.010 ^{a,b} (.0025)	.012 ^{a,b} (.003)	.011 ^{a,b} (.002)	6 ^{a,b} (0)	11 ^{a,b} (2)	2.1 ^{a,b} (.4)	.161 ^{a,b} (.026)
Household income per adult male equivalent squared ^a	-.06 ^{a,b} (.02)	-.003 ^{a,b} (.001)	-.17 (.13)	-.004 ^{a,b} (.002)	-.00004 ^{a,b} (.000015)	-.00005 ^{a,b} (.00002)	-.00006 ^{a,b} (.00002)	-.03 ^{a,b} (.01)	-.06 ^{a,b} (.02)	-.009 ^{a,b} (.003)	-.0008 ^{a,b} (.0002)
Weekly subsidy value of school lunches per adult male equivalent ^a	64 ^a (27)	3.75 ^{a,b} (.72)	78 (129)	6.71 ^{a,b} (1.52)	.073 ^{a,b} (.019)	.060 ^{a,b} (.022)	.102 ^{a,b} (.017)	7 (11)	20 (16)	9.7 ^{a,b} (3.1)	.570 ^{a,b} (.206)
Weekly subsidy value of school breakfasts per adult male equivalent ^a	122 (77)	3.37 (2.10)	292 (372)	3.42 (4.52)	.078 (.053)	.080 (.063)	.075 (.049)	61 (32)	109 ^a (45)	12.6 (8.9)	.467 (.626)
Weekly value of home-grown food per adult male equivalent ^a	195 ^{a,b} (21)	6.39 ^{a,b} (.69)	1,105 ^{a,b} (119)	8.87 ^{a,b} (1.45)	.122 ^{a,b} (.015)	.152 ^{a,b} (.018)	.128 ^{a,b} (.016)	70 ^{a,b} (8)	139 ^{a,b} (12)	26.4 ^{a,b} (2.8)	1.399 ^{a,b} (.145)
Weekly value of gift/pay food per adult male equivalent ^a	83 ^{a,b} (15)	3.23 ^{a,b} (.48)	343 ^{a,b} (83)	5.53 ^{a,b} (1.05)	.056 ^{a,b} (.011)	.070 ^{a,b} (.013)	.076 ^{a,b} (.01)	30 ^{a,b} (6)	49 ^{a,b} (9)	10.6 ^{a,b} (2.0)	.557 ^{a,b} (.108)
Female head present	-59 (134)	-2.31 (4.01)	2,124 ^{a,b} (680)	11.55 (7.21)	-.018 (.093)	.054 (.108)	-.138 (.085)	-160 ^{a,b} (39)	-279 ^{a,b} (56)	-16.2 (15.5)	-4.195 ^{a,b} (.621)
Black	-41 (74)	5.28 ^a (2.20)	3,763 ^{a,b} (373)	24.53 ^{a,b} (3.94)	-.019 (.051)	-.196 ^{a,b} (.108)	.103 ^{a,b} (.059)	-167 ^{a,b} (21)	-143 ^{a,b} (31)	-47.0 ^{a,b} (8.5)	.776 ^{a,b} (.335)
Number of adult-male-equivalent persons in household ^a	-140 ^{a,b} (26)	-4.13 ^{a,b} (.81)	-938 ^{a,b} (135)	-9.98 ^{a,b} (1.24)	-.073 ^{a,b} (.018)	-.090 ^{a,b} (.021)	-.090 ^{a,b} (.017)	-15 ^{a,b} (5)	-37 ^{a,b} (8)	-13.4 ^{a,b} (2.9)	-.403 ^{a,b} (.080)
Number of guest meals per adult male equivalent ^a	29 (16)	1.55 ^{a,b} (.48)	147 (83)	1.66 (1.04)	.013 (.011)	.026 ^a (.013)	.023 ^a (.011)	9 (6)	30 ^{a,b} (9)	4.4 ^a (2.0)	.354 ^{a,b} (.112)
North central	-23 (134)	-7.49 (4.01)	-1,069 (681)	-16.30 ^a (7.20)	.088 (.094)	-.042 (.108)	-.194 ^a (.085)	-21 (39)	-57 (56)	-19.0 (15.5)	.960 (.613)
South	366 ^{a,b} (113)	-7.30 ^{a,b} (3.38)	-1,182 ^{a,b} (574)	-16.70 ^{a,b} (6.07)	.324 ^{a,b} (.079)	.049 (.091)	-.168 ^a (.072)	63 (33)	136 ^{a,b} (47)	-7.9 (13.1)	1.767 ^{a,b} (.515)
West	-60 (148)	-8.54 (4.43)	-620 (752)	-3.23 (7.95)	.099 (.103)	-.080 (.120)	-.130 (.094)	35 (43)	8 (62)	15.2 (17.2)	1.089 (.677)
Spanish	581 ^{a,b} (132)	17.27 ^{a,b} (3.95)	1,417 ^{a,b} (670)	37.00 ^{a,b} (7.08)	.473 ^{a,b} (.092)	.365 ^{a,b} (.106)	.354 ^{a,b} (.084)	39 (38)	111 ^{a,b} (55)	19.6 (15.3)	2.898 ^{a,b} (.603)
Suburban	-184 (96)	-4.45 (2.86)	-1,194 ^a (485)	-15.34 ^{a,b} (5.12)	-.151 ^a (.067)	-.112 (.077)	-.132 ^a (.061)	-40 (28)	-81 ^a (40)	-22.1 ^a (11.1)	-1.145 ^{a,b} (.437)
Nonmetropolitan	33 (76)	-3.56 (2.28)	-1,848 ^{a,b} (387)	-19.02 ^{a,b} (4.09)	.019 (.053)	-.041 (.061)	-.118 ^a (.049)	35 (22)	31 (32)	.13 (8.83)	-.068 (.348)
Head of household is 35 to 59 years old	179 ^a (78)	6.40 ^{a,b} (2.35)	1,465 ^{a,b} (339)	10.85 ^{a,b} (4.16)	.126 ^a (.054)	.110 (.063)	.007 (.050)	63 ^{a,b} (23)	133 ^{a,b} (33)	15.3 (9.1)	2.140 ^{a,b} (.357)
Head of household is 60 years old or over	5 (92)	-8.55 ^{a,b} (2.67)	1,496 ^{a,b} (453)	5.86 (4.81)	-.068 (.063)	-.096 (.073)	-.268 ^a (.057)	34 (26)	4 (38)	-23.6 ^a (10.3)	2.230 ^{a,b} (.446)
R ²	.12	.12	.13	.10	.09	.09	.12	.12	.16	.11	.25
Mean of dependent variable	3,998 Kcal	128.57 g	11,414 IU	139.22 mg	2.715 mg	3.231 mg	2.560 mg	1,000 mg	1,710 mg	464 mg	17 mg

Source: 1979-80 Survey of Food Consumption in Low-Income Households, USDA.
 Note: The dependent variables are daily availability per equivalent nutrition unit (number of equivalent adult males eating from home food supplies). Equivalent nutrition units are computed separately for each nutrient and are based on the 1980 RDA.
^a The number of adult male equivalents is computed separately for each nutrient and are based on the 1980 RDA.
^b Single (double) asterisk(s) indicate statistically significant at the .05 (.01) level.

Table 3. Marginal Propensities to Consume Nutrients and Implied Total Effects

	Absolute MPC		MPC as a Percentage of the Adult Male RDA		Total Effects ^a	
	Food Stamp Benefit	Cash ^a	Food Stamp Benefit	Cash ^a	Food Stamp Benefit	Cash
Food energy (Kcal)	52.0	11.0%	1.9	.4	564	530
Protein (g)	1.81	.36	3.2	.6	19	18
Vitamin A (IU)	156	25	3.1	.5	1,691	1,256
Vitamin C (mg)	1.97	.59	3.3	1.0	21	29
Thiamin (mg)	.040	.006	2.9	.4	.434	.328
Riboflavin (mg)	.052	.008	3.3	.5	.564	.390
Vitamin B ₆ (mg)	.039	.007	1.8	.3	.423	.336
Calcium (mg)	18	4	2.3	.5	195	187
Phosphorus (mg)	30	7	3.8	.9	325	336
Magnesium (mg)	7.1	1.4	2.0	.4	77	68
Iron (mg)	.387	.115	3.9	1.2	4.195	5.038

Source: 1979–80 Survey of Food Consumption in Low-Income Households, USDA.

Notes: Absolute MPC is change in nutrient availability per ENU due to a one-dollar change in income per AME. Percentage MPC is absolute MPC divided by 1980 adult male RDA for the particular nutrient.

^a Defined as the level of nutrient availability attributed to food stamp benefits and cash income and calculated by multiplying the estimated coefficients by the average value of FSP benefits (for participants only) and cash income per AME.

food stamp benefit and cash income. These “total effects” are calculated by multiplying the estimated coefficients for the food stamp benefit, cash income, and cash income squared (all scaled by AME) by their respective average values for the benefit level and cash income for program participants. Except for vitamin C, phosphorous, and iron, the total availability of nutrients attributed to FSP benefits is either very close to or exceeds the availability of nutrients resulting from cash income.

Returning to table 2, the results indicate positive effects on nutrient availability of the subsidy value of school lunches and school breakfasts, although the latter is usually not statistically significant. The weekly value of home-grown and gift/pay food are also positively and significantly associated with nutrient availability and, in fact, the estimated coefficients for these two variables are significantly larger than the estimated coefficients for both the food stamp benefit and cash income variables. The number of AMEs in the household lowers nutrient availability per ENU. Nutrient patterns vary across region as well as by suburban and metropolitan residence, the two stratification variables used in the survey design.

Comparison of Results with Previous Findings

The OLS results indicate large and statistically significant differences between the estimated

coefficients of the food stamp benefit and cash income in the nutrient availability equations. The ratio of the MPC for the food stamp benefit to the MPC for cash income ranges from a low of 3 to a high of 7. The magnitude of these differences prompted us to examine more carefully the comparable results from previous studies to see if our findings are consistent with those in the literature.

Two studies were selected for comparison purposes, a study by Allen and Gadson and one by Basiotis et al. These two studies were selected because (a) both used data sources and had dietary outcome variables comparable to those used in our analysis, and (b) one had a model specification similar to that estimated in this study (Allen and Gadson) and the other a very different model specification (Basiotis et al.).⁶ Because these two studies used different model specifications, estimation procedures, and units of measurement (weekly versus daily measures of nutrient availability), their basic empirical results were adjusted to make their findings comparable. Because of the difference in model specification, the MPCs for the food stamp benefit could not be estimated from the Basiotis et al. study, but total impacts of the FSP could be.

The results from our study generally are sim-

⁶ In particular, the Basiotis et al. study included food expenditures as an independent variable in the nutrient availability equations and included the food stamp benefit as part of the household income variable.

ilar to those of Allen and Gadson. In particular, both studies report higher MPCs for the food stamp benefit than for cash income, and the magnitudes of the estimated coefficients are roughly the same across the two studies (with some exceptions). In addition, the estimated MPCs for cash income from the Basiotis et al. study are comparable in magnitude to those from this study.

However, estimates of the total effects of the FSP from this study generally are considerably larger than the comparable estimates from the Basiotis et al. analysis. This difference likely results from a model specification in that study in which the household income variable included the food stamp benefit, thus constraining the marginal effects of cash income and the food stamp benefit to be the same. Because the results of our study suggest differential effects of cash income and the food stamp benefit on nutrient availability, with larger MPCs for the food stamp benefit, the inclusion of the food stamp benefit in an overall household income measure leads to underestimates of the dietary effects of FSP benefits.

Selection Bias Models

The NFCS data used in the OLS regressions presented above include information on the food use of both FSP participants and FSP-eligible households not receiving FSP benefits (eligible nonparticipants). Because FSP participants are a self-selected group of low-income households, selection bias may exist if FSP households are high or low nutrient availability households to begin with, even before participation in the FSP and even if they had not participated in the FSP. The fact that eligible nonparticipants choose not to participate in the FSP suggests that they may differ systematically from participants in ways that may influence food consumption and, hence, nutrient availability. For example, if participating households are most in need of food assistance because of low initial food consumption, then OLS estimates of the effects of the food stamp benefit are downwardly biased. Conversely, participating households may have greater preferences for food than do eligible nonparticipants with similar household characteristics, leading to an upward bias in the OLS estimate of the effects of food stamps on nutrient availability.

Two types of selection bias are considered here, called Type A and Type B (Bjorklund and

Moffitt; and Ranney). The more common type of bias, Type A, arises when participants and nonparticipants in the FSP have different levels of food consumption, holding constant all other observed characteristics, even prior to participation in the FSP. In this type of selectivity, the problem arises because FSP participants and eligible nonparticipants differ, for whatever reason, in their initial level of food consumption. Type B selection bias arises if FSP participants and eligible nonparticipants have different marginal propensities to consume food (MPC) out of income. In this case, participants experience an increase in food consumption and nutrient availability that is different in magnitude than the increase that would be experienced by eligible nonparticipants if they were to participate.

The following set of equations represent Type A and Type B selection bias in model form:

$$(7) \quad N_{ki} = \alpha_k + \beta_{ki} (Y_i + \gamma_k B_i) X_i \phi_k + \epsilon_{ki}$$

$$(8) \quad P_i^* = Z_i \psi + \nu_i$$

$$(9) \quad P_i = 1 \text{ if } P_i^* \geq 0;$$

$$(10) \quad = 0 \text{ if } P_i^* < 0$$

$$\beta_{ki} = W_i \lambda_k + w_{ki},$$

where P_i^* is an index for the "propensity" to participate in the FSP, Z_i is a set of variables that affect that propensity, and P_i is a dummy variable equal to one if the household actually participates in the FSP and zero if not. Equations (8)–(9) represent a standard probit model for a dummy dependent variable. Included in Z_i is the potential food stamp benefit and cash income.⁷ If the error terms ϵ_{ki} and ν_i are correlated, Type A selection bias occurs. If they are positively correlated, OLS estimates of (7) yield upwardly biased estimates of the effects of food stamp benefits on nutrient availability; if they are negatively correlated, the opposite occurs.

This model has a single income variable, $Y + \gamma B$, where γ is the ratio of the MPC for the food stamp benefit to the MPC for cash income. The coefficient on this income variable, β_{ki} , is the MPC for income in general; it has a subscript "i" to allow for differences among households.

⁷ Specifically, the variables used as predictors of the likelihood of participating in the FSP are weekly cash income; potential food stamp benefit; race of the household head (1 = black, 0 = non-black); dummy variables for male head only or a female head only; dummy variables for the age, education, and employment status of the female household head (or male household head if there is no female household head); and a dummy variable for whether the household owns its home.

Thus, this model allows the MPC out of income to differ between FSP participants and eligible nonparticipants. As shown in equation (10), the MPC for income is assumed to be a function of a set of variables denoted by W and by an unobserved error term, w_i .⁸ Type B selection bias arises if the error terms w_i and v_i are correlated. If they are positively correlated, those households who have high MPCs even in the absence of the FSP (high w_i) are more likely to participate in the FSP.

Models with Type A selection bias and Type A and B selection bias were estimated using maximum likelihood estimation techniques (Maddala). That is, an FSP participation equation was estimated jointly with the nutrient

equations, using a maximum likelihood estimation technique and allowing the error terms of the participation and nutrient equations to be correlated.⁹ Given the complexity of selection bias models, the nutrient equations were estimated for only five of the nutrients: food energy, vitamin A, vitamin B₆, calcium, and iron. These five nutrients have the lowest average availability levels relative to the RDA and the lowest percentage meeting the RDA.

Table 4 shows the results of the important parameter estimates of the selection bias models for the five dietary components.¹⁰ Focusing first on the estimates for the Type A only selection bias model, the evidence for selection bias is weak. Of the five estimates of the cross-equa-

⁸ The set of variables assumed to influence the MPC for cash income are race, household size in adult-male-equivalent persons, number of guest meals eaten from household food supplies, dummy variables for the age (<35, 35-59, 60+) of the female household head, and dummy variables for whether the household lives in the South or in a suburban location.

⁹ With Type B selection bias, the individual-specific MPC is also assumed to be distributed normally and correlated with the error term in the nutrient equation.

¹⁰ The full set of results is available upon request from the authors.

Table 4. Results for Selection Bias Models

	Food Energy	Vitamin A	Vitamin B ₆	Calcium	Iron
MPC: Food stamp benefit					
OLS estimate	52	156	.039	18	.387
Type A selection bias	59	197	.046	14	.471
Type A and B selection bias ^a	57	175	.052	13	.365
MPC: cash income^b					
OLS estimate	11	25	.007	4	.115
Type A selection bias	12	26	.007	4	.122
Type A and B selection bias					
FSP participants	17	50	.009	6	.188
Eligible nonparticipants	18	55	.009	5	.188
Ratio of food stamp benefit to cash-income MPC					
OLS estimate	4.73	6.24	5.57	4.50	3.37
Type A selection bias	4.92	7.58	6.57	3.50	3.86
Type A and B selection bias	3.13	3.39	5.77	2.43	1.97
Cross-equation correlation coefficient, Type A^c					
(standard error in parentheses)	-.041 (.034)	-.051 (.035)	-.069 ^d (.035)	.047 (.034)	-.062 (.035)
Cross-equation correlation coefficients, Type A and B (standard error in parentheses)					
Correlation coefficient 1 ^e	-.022 (.085)	-.011 (.096)	-.126 (.081)	.059 (.084)	-.010 (.073)
Correlation coefficient 2 ^f	-.022 (.116)	-.072 (.146)	.047 (.167)	.016 (.122)	-.011 (.089)

Source: 1979-80 Survey of Food Consumption in Low-Income Households, USDA.

Note: Full set of FIML results are available upon request from the authors.

^a Evaluated at mean MPC for FSP participants.

^b Evaluated at mean cash income per AME.

^c Correlation between error terms in nutrient equation and FSP participation equation.

^d Significant at the .05 level.

^e Correlation between error terms in nutrient equation and FSP participation equation.

^f Correlation between error terms in MPC equation and FSP participation equation.

tion correlation coefficients, only one is statistically significant (vitamin B₆), and that estimated correlation coefficient is not large. With the exception of calcium, the sign of the correlation coefficients is always negative, implying that FSP participants would have lower nutrient availability levels than nonparticipants in the absence of the FSP. As a result, for the four nutrients with negative cross-equation correlation estimates, the estimates of the MPCs for the food stamp benefit and the ratio of the food stamp benefit to cash-income MPC from the selection bias model exceed the OLS estimates. Despite the general lack of significance of the cross-correlation coefficients, the percentage change in the MPCs for the food stamp benefit from the OLS and selection bias models range from 14% for food energy to 26% for vitamin A.

Table 4 also presents the estimates from the Type A and Type B selection bias models. The correlation parameters are all statistically insignificant. The correlation parameters between the nutrient and FSP participation equations are, as before, generally negative (except calcium), but now are often smaller and less significant. The second type of correlation parameters are sometimes positive and sometimes negative, and are always low in statistical significance (much lower than those for Type A). As a result, the estimated MPCs for the food stamp benefit and for cash income from the selection bias models are quite similar to the OLS estimates. In addition, the OLS finding that the MPC for the food stamp benefit is significantly larger than the cash-income MPC persists with the Type A and B selection bias model.

Summary and Conclusions

The major finding of the empirical analysis is that the estimated dietary effects of changes in FSP benefits are considerably larger than those resulting from changes in cash income. While both the food stamp benefit and cash income have significant positive effects on nutrient availability, the estimated effects of changes in the food stamp benefit on nutrient levels consistently exceed the estimated effects of changes in cash income. The ratios of the MPC for the food stamp benefit to the cash-income MPC are consistently and significantly greater than one. The OLS estimates of these ratios range from 3 to 7, and the estimates from the selection bias models for selected nutrients range from roughly 2 to 7.

These large differences between the estimated effects of food stamp benefit levels and cash income on nutrient availability are generally contrary to the implications of theoretical household expenditure models (Southworth, Senauer and Young), and it is important to determine if these differences truly reflect the relative efficiency of coupon versus cash benefits in enhancing the quality of diets of participating households. To do this, two econometric models of selection bias were estimated for this study. Type A selection bias tests whether households with different initial levels of nutrient availability are more or less likely to be FSP participants; Type B selection bias tests whether the change in nutrient availability per dollar of cash income is greater or smaller for FSP participants relative to eligible nonparticipants. The results of the selection bias models show little evidence of either type of selection bias. Moreover, the estimated dietary effects of the food stamp benefit and cash income from the selection bias models are very similar to those from the basic model estimated by OLS regression. This absence of significant selection bias is an important finding; it suggests that underlying differences in either preferences for food or food demand functions between FSP participants and nonparticipants cannot account for the large and statistically significant differences between the effects of food stamps and cash income on household nutrient availability.

[Received August 1989; final revision received February 1990.]

References

- Allen, J. E., and K. E. Gadson. "Nutrient Consumption Patterns of Low-Income Households." Washington DC: U.S. Department of Agriculture, Econ. Res. Serv. Tech. Bull. No. 1685, 1983.
- Basiotis, P., M. Brown, S. Johnson, and K. Morgan. "Nutrient Availability, Food Costs, and Food Stamps." *Amer. J. Agr. Econ.* 65(1983):685-93.
- Bjorklund, A., and R. Moffitt. "The Estimation of Wage Gains and Welfare Gains in Self-Selection Models." *Rev. Econ. and Statist.* 69(1987):42-49.
- Chen, Jain-Shing A. "Simultaneous Equations Models with Qualitative Dependent Variables: A Food Stamp Program Participation and Food Cost Analysis." Ph.D. thesis, University of Missouri, 1983.
- Johnson, S., J. Burt, and K. Morgan. "The Food Stamp Program: Participation, Food Cost, and Diet Quality for Low Income Households." *Food Tech.* 35(1981):58-70.
- Maddala, G. S. *Econometrics*. New York: McGraw-Hill Book Co., 1977.

- Ranney, C. "A Study of the Interdependent Food Stamp Program Participation and Food Demand Decisions." Ph.D. thesis, University of California, Davis 1983.
- Senauer, B., and N. Young. "The Impact of Food Stamps on Food Expenditures: Rejection of the Traditional Model." *Amer. J. Agr. Econ.* 68(1986):37-43.
- Southworth, H. M. "The Economics of Public Measures to Subsidize Food Consumption." *J. Farm Econ.* 27(1945):38-66.
- Smallwood, D. M., and J. R. Blaylock. "Analysis of Food Stamp Program Participation and Food Expenditures." *West. J. Agr. Econ.* 10(1985):41-54.

Demand for Food Fats and Oils: The Role of Demographic Variables and Government Donations

Brian W. Gould, Thomas L. Cox, and Federico Perali

A systems model of U.S. food fats and oils demand is estimated using quarterly time-series data for the period 1962–87. Demographic scaling is used to incorporate demographic variables and per capita government butter donations. In addition to price, income, and demographic demand elasticities, dietary fat intake elasticities are calculated for each of the demographic characteristics included in the study. All of the five own-price elasticities are statistically significant, as are fifteen of the twenty cross-price elasticities. Fourteen of the fifteen demographic elasticities are statistically significant.

Key words: AIDS, demand system, demographic scaling, nutrient elasticities.

Increased public concern with fat consumption, especially saturated fat and blood cholesterol levels, is evidenced by the 1988 report of the Surgeon General (U.S. Department of Health and Human Services). Over the 1962–85 period, total dietary fat consumption increased from 143 grams per day to 169 grams per day, an 18% increase. Over this same period, the per capita consumption of saturated fat remained relatively constant (58–60 grams/day), while polyunsaturated fat consumption increased from 22 to 33 grams per day (i.e., a 50% increase) and mono-unsaturated fat consumption increased from 59 to 68 grams per day (i.e., a 15% increase, Raper and Marston).

Red meats and food fats and oils are the two major sources of dietary fat. Both total and unsaturated fat intake from food fats and oils has increased relative to the fat obtained from red meat sources. This increase reflects the increased consumption of poultry and fish in place

of red meats, the increased consumption of food fats and oils, and the fat characteristics of these foods. Over the 1957–59 period, 40.9% and 54.5% of total dietary and polyunsaturated fat, respectively, were obtained from food fats and oils. These contributions increased to 47.2% and 68.2% by 1985 (Raper and Marston).

Figure 1 shows U.S. per capita disappearance of the five food fats and oils included in this study: butter, margarine, vegetable shortenings, salad and cooking oils, and lard. Steady increases in the consumption of vegetable shortenings, salad and cooking oils, lower consumption levels of animal-based fats and oils (i.e., butter and lard), and relatively constant margarine disappearance are shown.

The relative changes in the consumption of these food fats and oils raises questions about the causes. In this study a demand system is used to identify and measure the impacts of such factors. The specific objectives are threefold: (a) to incorporate a set of demographic variables into a demand systems model in a manner which avoids the shortcomings of previous efforts; (b) to examine how the per capita demand for food fats and oils is affected by a changing demographic profile (age, education, racial composition) and government butter donations; and (c) to examine how the dietary intake of total, saturated, and unsaturated fat has been affected by changes in the demographic profile of the U.S. population.

The authors are, respectively, an associate scientist, Center for Dairy Research and Department of Agricultural Economics; an assistant professor and a graduate research assistant, Department of Agricultural Economics, University of Wisconsin. Seniority of authorship is equally shared.

Financial support provided by the Wisconsin Milk Marketing Board and the College of Agriculture and Life Sciences, University of Wisconsin, is acknowledged.

The authors thank Rueben Buse and Jean-Paul Chavas for helpful comments on an earlier draft of this manuscript. Any remaining errors or omissions remain the responsibility of the authors.

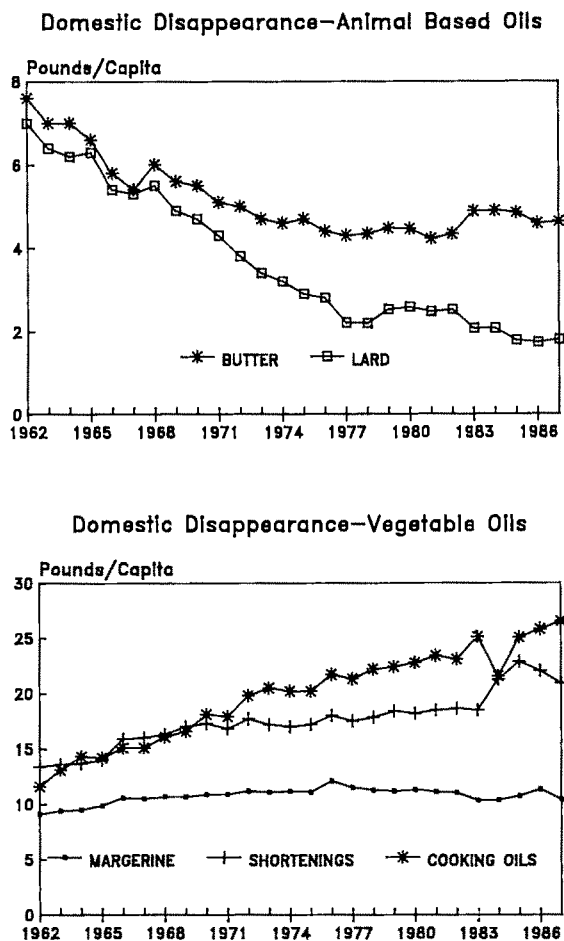


Figure 1. Per capita disappearance of food fats and oils, 1962–87

Description of the Demand System

Food choices are affected by demographic, social, and economic factors. In the model developed here, we hypothesize that the U.S. consumption of food fats and oils is determined by relative prices and tastes and preferences which, in turn, are reflected by socioeconomic variables. A set of demographic variables is incorporated into a demand system in which the demand for fats and oils is separable from the demand for other foods. That is, the consumer first decides on the level of fat and oils expenditure and then allocates these expenditures between the various food fats and oils sources based on relative prices and demographic characteristics.

Demographic characteristics can be modeled in several ways. The use of equivalence scales is one method; however, with this approach de-

mographic characteristics are virtually equivalent to changes in prices (Lewbel, p. 1). This restriction has been overcome through demographic translating, scaling, and Gorman procedures. In this study the demographic scaling procedure originally proposed by Barton is used so that the demographic variables can affect more than the "subsistence" or "necessary" parameters of the demand system (Pollak and Wales 1981).¹

Representing the original demand system as

$$(1) \quad D_i = D_i(P, S, M) \quad (i = 1, \dots, n),$$

where D_i is the per capita demand for the i th commodity, P is a vector of commodity prices, S is a vector of demographic characteristics, M is a given level of expenditure, and n is the number of commodities. Demographic scaling modifies this system as

$$(2) \quad \begin{aligned} D_i(P, S, M) &= \phi_i D_i^*(p_1 \phi_1, p_2 \phi_2, \dots, p_n \phi_n, M) \\ &= \phi_i D_i^*(p_1^*, p_2^*, \dots, p_n^*, M), \end{aligned}$$

where $p_i^* = p_i \phi_i$ are scaled prices and the ϕ_i 's are scaling parameters which are functions of the demographic variables, s_1, \dots, s_d :

$$(3) \quad \phi_i = \phi_i(s_1, \dots, s_d).$$

The scaling functions represent the number of commodity-specific "profile equivalents" (Pollak and Wales 1980). With respect to the demand for food fats and oils, preferences depend on the number of pounds per profile equivalent, i.e., pounds/ ϕ_i , and on the price per pound per profile equivalent, $p_i \phi_i$, where i is butter, margarine, etc., and ϕ_i is a profile of the consumer group, which is a function of demographic characteristics (Pollak and Wales 1981, pp. 1535–36).

Empirical analysis of the demand for food fats and oils using this model requires a functional form for (2) and (3). The almost ideal demand system (AIDS) model originally formulated by Deaton and Muellbauer is adopted because of the lack of a priori restrictions imposed on the substitution characteristics between commodities and the exact aggregation properties of this functional form (Deaton and Muellbauer). A

¹ Ray (1982) and Rossi apply the scaling function to household-level data. As noted by Rossi and by Deaton and Muellbauer, the scaling function can be interpreted as a measure of household size that takes into account age composition, other household characteristics, and economies of household size which can be used to deflate total expenditures to reflect a "needs corrected per capita level" (Deaton and Muellbauer, p. 314).

scaling function similar to the specification used by Ray (1982) and by Barnes and Gillingham is used to estimate equation (4):

$$(4) \quad \phi_i = \prod_r s_r^{\epsilon_{ir}},$$

where ϵ_{ir} are estimated coefficients.² Incorporating this scaling function into the AIDS model yields

$$(5) \quad w_i = \alpha_i + \sum_j \delta_{ij} \ln p_j^* + \beta_i \ln(M/P^*)$$

$$(i = 1, \dots, n),$$

where w_i is the share of the i th type of food fat and oil, and

$$(6) \quad \ln P^* = \alpha_0 + \sum_k \alpha_k \ln p_k^* + \frac{1}{2} \sum_k \sum_j \delta_{kj} \ln p_k^* \ln p_j^*.$$

With the incorporation of the scaling functions, the traditional homogeneity and adding up conditions still hold.³ In addition, equation (5) must be homogenous of degree zero in p_i^* as well as p_j . In order to maintain homogeneity with respect to p_i , (5) must be homogenous of degree zero with respect to the demographic variables. That is,

$$(7) \quad \sum_i \epsilon_{ir} = 0 \quad (r = 1, \dots, d).$$

Similar to previous studies that have incorporated demographic variables in demand systems, a linearized index (such as Stone's index using scaled prices) is assumed to provide an acceptable approximation to $\ln P^*$ (Capps, Tedford, and Havlicek). As such, equation (6) is approximated by

$$(8) \quad \ln P^* \approx \sum_k w_k \ln p_k^*.$$

In order to avoid simultaneity problems, we adopt

the Eales and Unnevehr approach where lagged budget shares are used in equation (8).⁴

In the AIDS model, expenditure, price, and demographic elasticities change over time as the share of each commodity changes. Differentiating equation (5), the expenditure (N_i), uncompensated price elasticities (Γ_{ij}), and compensated price elasticities (Γ_{ij}^*) can be calculated as

$$(9) \quad N_i = 1 + \beta_i / w_i,$$

$$(10) \quad \Gamma_{ij} = ((\delta_{ij} - \beta_i w_{j,t-1}) / w_i) - K_{ij},$$

$$(11) \quad \Gamma_{ij}^* = ((\delta_{ij} - \beta_i w_{j,t-1}) / w_i) - K_{ij} + w_j [1 + \beta_j / w_j],$$

where K_{ij} is equal to 1 if $i = j$, 0 otherwise. Similarly, differentiating (2) with respect to the r th demographic characteristic and converting to elasticity format (Pollak and Wales 1980), the elasticity of the demand for the i th good with respect to the r th demographic characteristic (E_{ir}) is a function of the uncompensated own- and cross-price elasticities (Γ_{ij}) and the elasticity of the i th goods scaling function with respect to the r th demographic characteristic, Ω_{ir} :

$$(12) \quad E_{ir} = \Omega_{ir} + \sum_j \Gamma_{ij} \Omega_{jr}.$$

Description of the Data

The demand for the five types of food fats and oils are analyzed using aggregate quarterly U.S. time-series data for 1962–87.⁵ Commercial butter disappearance data are obtained from various issues of *Dairy Situation* (USDA). Quarterly commercial disappearance for margarine, vegetable shortening, and salad and cooking oils are estimated from the following:

$$(13) \quad \text{COMMDIS}_t = \text{PROD}_t + \text{STOCK}_{t-1} + \text{IMPORT}_t - \text{EXPORT}_t - \text{STOCK}_t,$$

where *COMMDIS* is commercial disappearance, *PROD* represents production levels, *STOCK* are ending stock levels, *IMPORT* are imports, and *EXPORT* represents exports of the commodity. Various issues of *Current Industrial Reports: Fats and Oils Production, Consumption, and Stocks*

² As noted by an anonymous reviewer, the estimated scaling coefficients are assumed to remain constant over the study period. This assumption implies that, as more health information becomes available, the role of demographic characteristics in determining per capita consumption is unaffected. Future research could develop a time-varying scaling specification similar to the dynamic AIDS models developed by Moschini and Meilke and by Ray (1984).

³ The usual symmetry and homogeneity conditions on the AIDS share equations are adopted. We conducted log-likelihood ratio tests (Judge et al., p. 758) and could not reject these restrictions.

⁴ p^* is used in place of p in Stone's index because the underlying cost function is now dependent on p^* .

⁵ One problem with the use of this aggregate per capita disappearance data is that they encompass both intermediate and final commodity demand. Thus, the econometric results should be interpreted with caution.

are used to obtain monthly production and stocks (U.S. Department of Commerce). Annual import and export data are obtained from Putnam. For lard, various issues of the *CRB Commodity Yearbook* are used to obtain monthly commercial disappearance values (Buchman and Gaylann). Estimates of aggregate per capita disappearance values are obtained by dividing by quarterly population estimates obtained from published Department of Commerce sources.

Quarterly estimates of butter, margarine, and shortening prices are obtained from published BLS retail city average price series. Wholesale prices for salad and cooking oils and lard are converted to retail prices using data provided in various issues of the *Fats and Oils Situation* (USDA). These data, as well as those used for estimation, are available upon request from the authors.

The demographic variables are as follows: the median age of the population (*AGE*), the non-white proportion of the population (*NON-WHITE*), and the median years of schooling completed by those over twenty-five years of age (*SCHOOL*). These data are obtained from Bureau of the Census (*Current Population Reports* and *Statistical Abstract of the United States*). Quarterly values of the demographic variables are estimated by interpolating from the annual estimates of these variables. The variable *DO-NATE*, which represents per capita USDA donations of butter, is also used as an explanatory variable in each of the commodity-specific scaling functions. Estimates of quarterly government donations are obtained from *Dairy Situation* and other published USDA data (*CCC Milk Price Support and Related Activities*).⁶

Total expenditure for fats and oils (*EXPEND*) is the sum of the expenditures of each of the component commodities. Commodity-specific expenditure values and budget shares are obtained from the above estimates of prices and quantities.

Estimation Procedure

Given the adding-up characteristics of the AIDS share equations, only four equations in a five-

equation system are independent. Thus, the lard equation is dropped from the estimation, with the parameters of this equation estimated from the symmetry and homogeneity conditions. Because the model with demographic scaling is nonlinear in the parameters, an iterative seemingly unrelated regression procedure (SYSNLIN within the SAS system) is used for estimation. Price, expenditure, demographic, and donation elasticities are calculated from the estimated coefficients and corresponding share values. Approximate asymptotic standard errors are computed via the δ -method for the demographic elasticities because of the nonlinear nature of the elasticities (Rao).

The effects of seasonality in food fats and oils consumption are accounted for by several harmonic variables in each share equation.⁷ Following Doran and Quilkey, and Kinnucan and Fearon, these harmonic variables are calculated as

$$(14) \quad H_{1kt} = \cos(\lambda_k t) \quad (k = 1, 2) \\ (t = 1, \dots, T), \\ H_{2kt} = \sin(\lambda_k t) \quad (t = 1, \dots, T),$$

where

$$(15) \quad \lambda_k = 2\pi k/4 \quad (k = 1, 2).$$

With the incorporation of the scaling function [equation (4)] and the harmonic variables [equation (14)], the estimated AIDS share equations are

$$(16) \quad w_i = \alpha_i + \sum_j \delta_{ij} \left(\ln p_j + \sum_r \epsilon_r \ln s_r \right) \\ + \beta_i \left(\ln M - \left(\sum_j w_{j,t-1} \left(\ln p_j + \sum_r \epsilon_r \ln s_r \right) \right) \right) \\ + h_{i1} H_{11} + h_{i2} H_{21} + h_{i3} H_{12} \quad (i = 1, \dots, 5).$$

Estimated Coefficients and Elasticities

Table 1 presents the estimated demand system coefficients and associated asymptotic standard errors. All own-price coefficients are statistically significant at the 5% level except for the own price of lard. Six of the ten cross-price

⁶ Government donations are calculated as follows:

$$GD_t = GOVSTOCK_{t-1} - GOVSTOCK_t \\ + REMOVE_t - FDONATE_t,$$

where *GD* is government donations, *GOVSTOCK* is government ending stocks, *REMOVE* represents government net removals and *FDONATE* is foreign donation commitments.

⁷ A dummy variable specification was also evaluated to address seasonality in U.S. fats and oils consumption. Demand system parameters (price, conditional expenditure, and demographic effects) were quite robust across the dummy variable and harmonic specifications. Results for the harmonic specification are presented because it provided marginally superior explanatory power in terms of higher R^2 and log-likelihood values.

Table 1. AIDS Share-Equation Estimates for U.S. Fats and Oils Sector Using Quarterly Aggregate Disappearance Data (1962-87)

Dependent Variable	Independent Variables														R ²
	Price						Harmonic Variables								
	Butter	Margarine	Short	Cooking	Lard	Expenditure	COS1	COS2	SIN1	AGE	SCHOOL	NONWHITE	DONATE	Intercept	
Butter	.0801 (.0109)	-.0219 (.0063)	-.0238 (.0103)	-.0270 (.0082)	-.0074 (.0254)	.0209 (.0097)	.0093 (.0022)	.0058 (.0015)	.0002 (.0024)	5.5383 (1.2449)	-5.1744 (1.2212)	-2.7120 (.7213)	-.0055 (.0043)	.4065 (.2772)	.915
Margarine		.1321 (.0174)	-.0289 (.0138)	-.0643 (.0052)	-.0170 (.0332)	.0010 (.0026)	.0064 (.0010)	-.0002 (.0007)	.0078 (.0010)	-3.6836 (.9038)	-2.396 (.7242)	2.0452 (.5382)	.0014 (.0015)	1.6593 (.3132)	.844
Shortening			.1280 (.0171)	-.0667 (.0105)	-.0086 (.0370)	-.0251 (.0104)	-.0025 (.0019)	-.0030 (.0013)	-.0018 (.0020)	.6708 (1.2442)	4.4931 (1.0309)	-.0201 (.6659)	.0057 (.0026)	-1.5213 (.4305)	.944
Salad and cooking oils				.1723 (.0098)	-.0144 (.0241)	-.0056 (.0056)	-.0151 (.0024)	-.0033 (.0017)	-.0045 (.0025)	-3.3381 (.7672)	2.6732 (.7536)	3.0907 (.4436)	-.0008 (.0027)	-.0118 (.2538)	.913
Lard					.0473 (.0608)	.0088 (.0216)	.0018 (.0055)	.0007 (.0037)	-.0017 (.0058)	.8126 (2.7608)	-1.7523 (2.4926)	-2.4038 (1.5770)	.0008 (.0079)	.4673 (.7205)	

Source: Computed by authors.

Note: The model was estimated via iterative seemingly unrelated regression (ITSUR) using the SYSNLIN procedures of SAS. Asymptotic standard errors are presented in parentheses. The R²-statistic is computed as 1 - (sum of squared errors/corrected sum of squares of the actual dependent variable).

coefficients are also significant at this level, as is the butter and shortenings expenditure coefficients. Seven of the harmonic variable coefficients and nine of the estimated demographic coefficients are statistically significant at the 5% level.⁸

Table 2 shows the uncompensated price and expenditure elasticities evaluated at the means of the data. All of the own-price elasticities are statistically different from zero at the 5% level, as are fifteen of the twenty cross-price elasticities.⁹ In terms of the relationship between butter and margarine, the cross-price elasticity of the change in butter's price on margarine demand is statistically significant at the 5% level and is negative, indicating gross complementarity. Similar results are found by Goddard and Amuah, and by Cox for the Canadian fats and oils market. Similarly, margarine price has a statistically significant (gross) effect on butter (Marshallian) demand. All significant elasticities indicated gross complementarity except for lard and butter.

Table 3 provides a comparison of price and expenditure elasticities estimated here and those of Goddard and Amuah, Cox, and Huang. Goddard and Amuah, and Cox used Canadian quarterly data. To provide as close a comparison as possible, the elasticities obtained in the present study are reestimated using means of the independent variables over the 1973-86 period, the period over which Goddard and Amuah, and Cox estimated their models. The Canadian butter and margarine price and expenditure elasticities obtained by Goddard and Amuah, and by Cox are quite similar to those obtained in the current study. The results obtained by Huang are quite different; they may reflect the use of annual time-series data and a different time period for analysis.

From equation (12) we calculate the elasticity

⁸ The existence of fourth-order autocorrelation was examined by incorporating a variable representing residuals lagged four periods and estimated following the procedures proposed by Berndt and Savin. A log-likelihood ratio test of the null hypothesis that the autocorrelation coefficient is zero could not be rejected, indicating that fourth-order autocorrelation is not a problem (Judge et al., p. 758). With the harmonic variables, the values of the intercepts across quarters changed very little. For example, for butter the range was from .4004 in the third quarter to .4216 in the fourth quarter. The significance of the intercepts did not change appreciably across quarters.

⁹ Elasticities of substitution are estimated from the price elasticities presented in table 2. All of the own substitution elasticities are negative and significantly different from zero at the 5% level. Butter is a net substitute for margarine, vegetable shortenings, and cooking oils. Margarine is a net substitute for vegetable shortenings. None of the cross-substitution elasticities for lard are statistically significant.

Table 2. Uncompensated Price, Demographic, and Donation Elasticities Evaluated at the Sample Means (1962–87)

Dependent Variable	Price					Demographic Variables				
	Butter	Margarine	Short	Cooking	Lard	Expenditure	AGE	SCHOOL	NONWHITE	DONATE
Butter	-.662 (.057)	-.114 (.029)	-.128 (.045)	-.150 (.035)	.040 (.020)	1.094 (.040)	2.654 (.439)	-2.579 (.513)	-1.465 (.285)	-.003 (.002)
Margarine	-.130 (.037)	-.227 (.112)	-.171 (.082)	-.379 (.032)	-.100 (.022)	1.006 (.015)	-2.497 (.374)	-1.118 (.571)	1.008 (.221)	.001 (.001)
Shortening	-.079 (.048)	-.107 (.060)	-.421 (.096)	-.255 (.046)	-.030 (.022)	.891 (.049)	1.195 (.435)	2.346 (.524)	-.790 (.257)	.004 (.002)
Salad and cooking oils	-.083 (.027)	-.204 (.017)	-.210 (.035)	-.440 (.047)	-.045 (.014)	.982 (.019)	-1.751 (.260)	1.099 (.361)	1.642 (.126)	-.001 (.002)
Lard	-.146 (.069)	-.285 (.061)	-.164 (.095)	-.264 (.061)	-.277 (.112)	1.136 (.056)	1.574 (.748)	-1.812 (.851)	-2.667 (.583)	-.001 (.003)

Source: Computations by the authors.

Note: Because of the nonlinear nature of the elasticities, approximate asymptotic standard errors computed via the δ -method are provided in parenthesis.

of commodity demand with respect to changes in the three demographic characteristics (table 2). Fourteen of the fifteen demographic elasticities are statistically different from zero at the 5% level or better (i.e., all but *AGE/Lard*). *AGE*

has elastic and positive effects on butter, shortening, and lard demand and a negative impact on margarine and cooking oils. These results suggest the aging of the U.S. population increases (decreases) butter (margarine) demand.

Table 3. Comparison of Alternative Fats and Oils Price Elasticity Estimates for the U.S. and Canada Evaluated at the 1973–86 Data Means

Previous Study	Dependent Variable	Price Elasticity				Expenditure Elasticity
		Butter	Margarine	Shortening	Cooking Oils	
Butter						
U.S.						
Current		-.62	-.13	-.15	-.17	1.11
Huang ¹		-.17	.05	-.12		.02
Canada						
Goddard and Amuah		-.72	-.29	-.09	-.08	1.18
Cox		-.79	-.12	-.16	-.08	1.33
Margarine						
U.S.						
Current		-.13	-.23	-.17	-.37	1.01
Huang		.07	-.27	.19		.11
Canada						
Goddard and Amuah		-.26	-.60	-.04	-.07	.84
Cox		-.14	-.47	-.31	-.07	.92
Shortening						
U.S.						
Current		-.08	-.10	-.45	-.24	.90
Huang		-.03	.03	-.22		.37
Canada						
Goddard and Amuah		-1.19	-.57	.72	-.74	1.77
Cox		-.15	-.24	-.26	-.21	.53
Salad and cooking oils						
U.S.						
Current		-.08	-.19	-.20	-.48	.98
Goddard and Amuah		.03	.62	-.52	-.14	.01
Canada						
Cox		-.31	-.24	.02	-.57	.93

¹ Huang delineated only three categories: butter, margarine, and other oils. The other oils category is presented in the shortening rows and columns. The elasticities estimated by Huang were obtained from an annual time-series model of the years 1953–81.

ceteris paribus. That is, older Americans have stronger preferences for butter relative to margarine, perhaps reflecting preferences acquired prior to the recent negative health concerns for cholesterol and saturated fat in the diet.¹⁰

The highly elastic, statistically significant and negative (positive) *SCHOOL* elasticity in the butter and lard (shortening) equations indicates that increases in the average U.S. level of education have reduced (increased) butter and lard (shortening and cooking oils) demand, *ceteris paribus*. These results are intuitively appealing under the hypothesis that higher education levels increase dietary/health awareness and concern. The result is a *ceteris paribus* substitution away from lard (and butter) to vegetable shortenings. The sign and magnitude of the *NON-WHITE* elasticities indicate that increases in the proportion of the nonwhite population have decreased (increased) the demand for butter, shortening, and lard (margarine and cooking oils), *ceteris paribus*.

The estimated coefficients of the government donations variable (*DONATE*) can be used to analyze the impacts of government butter donations on the demand for the five commodities in the study. While government butter donations are found to have small negative elasticity impacts on commercial disappearance of butter (-.003), these impacts are not statistically different from zero. The associated shortening elasticity is statistically significant and positive, although relatively small (table 2). One explanation of the magnitude and significance of the *DONATE/Shortening* elasticity (0.004) may be that the income gained from government donations is reallocated to the consumption of shortenings, a net substitute for butter (table 4). Hence, as butter prices decrease for the (relatively small) recipient portion of the population (i.e., it is free to those who receive the donations, but they are a small part of the total population), the complementary price effect of butter on shortenings (-.079, table 2) and the positive conditional expenditure elasticity (.891, table 2) combine to yield a small but statistically significant positive impact on shortening demand. However, given the conditional nature of these second-stage estimates, caution in their interpretation is advised.

¹⁰ As the U.S. population ages, the proportional impact of this older cohort increases. Hence, the *AGE* variable likely reflects cohort rather than pure age effects.

Table 4. Fat Composition of Food Fats and Oils

Type of Food	Fat Content (Grams/Pound)				
	Total Fat	Saturated Fat	Unsaturated Fat		
			Mono	Poly	Total
Butter	370	229 (62)	106 (29)	14 (4)	120 (33)
Margarine	366	72 (20)	163 (45)	115 (31)	278 (76)
Shortenings	454	114 (25)	202 (45)	118 (26)	320 (81)
Salad and cooking	454	68 (15)	195 (43)	171 (38)	366 (81)
Lard	454	178 (39)	205 (45)	51 (11)	256 (56)

Source: Guthrie, appendix A.

Note: The salad and cooking oils coefficients pertain to soybean oil (partially hydrogenated). The shortening coefficients pertain to vegetable shortenings. The figures in parenthesis correspond to the percent of total fat. They do not sum to 100 due to the inability to correctly identify and measure all fatty acids.

Changing Demographic Characteristics and Fat Consumption

Given the assumed separability of the consumption of food fats and oils from other commodities and following Pitt and Sahn, the elasticity response of the intake of the *m*th-type of dietary fat (i.e., total, saturated, and unsaturated fat) obtained from food fats and oils to a change in the *r*th demographic characteristic (*D_{mr}*) can be calculated:

$$(17) \quad D_{mr} = \left(\sum_i a_{mi} E_{ir} Q_i \right) / \left(\sum_i a_{mi} Q_i \right),$$

where *a_{mi}* is the quantity of the *m*th-type of dietary fat per unit of the *i*th-type of food, *E_{ir}* is the demographic quantity elasticity calculated via (12), and *Q_i* the predicted quantity of the *i*th food. The values of *a_{mi}* are obtained from table 4. The "fat intake" elasticities estimated via equation (17) are conditional because of the separability assumption, i.e., the fat intake associated with meat or other dairy products is not covered by this analysis.

Table 5 presents these elasticity estimates (with associated standard errors) for the mean values of the independent variables over the entire study period as well as point estimates for three quarters. Approximate standard errors obtained by the δ -method are also computed to facilitate interpretation. The predicted quantities for each type of fat or oil are obtained by using the estimated shares and actual prices for the three pe-

Table 5. Estimated Demographic Nutrient Elasticities, 1962-87 and Second Quarter 1967, 1977, and 1987

Variable	Type of Fat	1962-87	Second Quarter		
			1967	1977	1987
<i>AGE</i>					
	Total fat	-.355 (.068)	-.220 (.065)	-.224 (.074)	-.688 (.094)
	Saturated fat	.423 (.064)	.522 (.085)	.568 (.060)	.074 (.058)
	Total unsaturated fat	-.637 (.103)	-.495 (.094)	-.478 (.102)	-.920 (.121)
<i>SCHOOL</i>					
	Total fat	.461 (.078)	.403 (.084)	.474 (.082)	.580 (.120)
	Saturated fat	-.090 (.088)	-.127 (.120)	-.060 (.088)	.077 (.085)
	Total unsaturated fat	.661 (.108)	.601 (.103)	.646 (.107)	.734 (.148)
<i>NONWHITE</i>					
	Total fat	.170 (.039)	.001 (.038)	.091 (.042)	.482 (.054)
	Saturated fat	-.446 (.040)	-.593 (.049)	-.511 (.036)	-.172 (.036)
	Total unsaturated fat	.391 (.058)	.220 (.055)	.282 (.058)	.679 (.069)

Source: Computations by the authors.

Note: Because of the nonlinear nature of the elasticities, approximate asymptotic standard errors computed via the δ -method are provided in parenthesis. These elasticities are conditional on total expenditures for food fats and oils. These elasticities refer to the dietary intake of fat from these commodities. Hence, these elasticities do not include the fat intake from meat and other dairy products.

riods covered by this table. When viewing table 5, it should be noted that the coefficients for margarine, vegetable shortening, and salad and cooking oils presented in table 4 are aggregates of several different types of oils with varying fat compositions. The results may differ with a finer definition of food fats and oils.

The total fat elasticity for *AGE* is negative and statistically significant when evaluated over the 1962-87 period and for the three quarters evaluated. This implies that the aging of the U.S. population is associated with a statistically significant, *ceteris paribus* reduction of total fat intake from the commodities included in the model. In terms of the composition of this total fat consumption, a positive (negative) and statistically significant relationship between *AGE* and saturated (unsaturated) fat intake is obtained. These results primarily reflect the elastic impacts of *AGE* on butter (+), margarine (-), shortening (+), cooking oils (-), and lard (+) from table 3 and the associated fat compositions of these commodities (table 1).

Total fat consumption from the fat and oil commodities is positively affected by increased education (*SCHOOL* elasticities, table 5). These

results suggest that increasing education levels have increased consumption of unsaturated fats more than the associated decrease in saturated fats. These results are appealing under the hypothesis that, with higher education levels, the population has a greater ability to obtain information on, and understand the health impacts of, the consumption of certain types of fats. Conversely, one might expect better educated consumers to decrease their total fat intake (i.e., following dietary guidelines), not just saturated fat intake. The positive unsaturated fat *SCHOOL* elasticities may also reflect increased purchases of food away from home (*FAFH*) for those individuals with higher income levels. Much of these *FAFH* purchases occur in fast food establishments that use a large amount of vegetable-based fats.

The changing ethnic composition of the U.S. population appeared to have little impact on the intake of total fat except for the latter part of the 1980s. In contrast, the saturated fat elasticities were negative and statistically different from zero (5% level) for the late 1960s and 1970s. The unsaturated fat elasticities are positive and significant at all points evaluated.

Concluding Comments

This research analyzes the impacts of prices, fats and oils expenditures, seasonality, government butter donations, and demographic factors on quarterly U.S. fats and oils consumption over the 1962–87 period. The demand systems approach using an AIDS with demographic scaling specification performs quite well statistically and in terms of the implied elasticity results. In particular, the butter and margarine Marshallian price elasticities are similar (in signs and magnitudes) to the results of Goddard and Amuah, and Cox. The estimated impacts of the changing U.S. demographic profile [in particular, the aging of the population (*AGE*), increased average levels of schooling (*SCHOOL*), and increased proportion of the population that is nonwhite (*NON-WHITE*)] on fats and oils consumption are generally statistically significant and appear intuitively reasonable.

The nutrient/demographic elasticities indicate that the aging of the population (*AGE*) is associated with statistically significant decreases in total fat intake, but with increases in the amount of saturated fats and decreases in the amounts of unsaturated fats. One interpretation of these *AGE* results is that older consumers acquired strong (positive) preferences for butter (and vegetable shortening) prior to current concerns for cholesterol and saturated fat in the U.S. diet.

Increasing average education level (*SCHOOL*) is associated with statistically significant increases in total fat intake, but with negative (positive) impacts on the amount of saturated (unsaturated) fat from the commodities modeled. Increases in the proportion of the nonwhite population (*NONWHITE*) have little impact on the total intake of fats but to increase (decrease) the intake of unsaturated (saturated) fats. These results are conditional in so far as the fat consumption from sources other than those modeled here (e.g., meats and other dairy products) are not reflected in these results.

Government donations (*DONATE*) of butter have statistically insignificant and small negative impacts on commercial disappearance of butter. The results suggest that the potential reallocation of the "income transfer" generated by these donations is spent on shortenings; but, given the conditional (second stage) nature of the specification estimated, this result should be interpreted cautiously.

The inclusion of demographic information in the estimation of conditional demand systems

from aggregate disappearance data is strongly supported by this research. These procedures explicitly model changes in tastes and preferences associated with changes in the demographic profile, allow estimation of profile specific response parameters, and provide a natural basis for projecting demand based on demographic projections.

The success obtained here with respect to the use of demographic characteristics in demand systems estimation supports the continuation of research first initiated by Barton. Three areas of research that could extend the model developed here are (a) development of a model where the demographic coefficients are allowed to vary over time, (b) application of the model to household level cross-sectional data, and (c) extension of the Pollack and Wales (1981) demand systems model where the "appropriate" method (i.e., translating, scaling, Gorman) for incorporating demographic effects is parametrically tested when multiple demographic characteristics are included in the model.

[Received January 1990; final revision received April 1990.]

References

- Barnes, R., and R. Gillingham. "Demographic Effects in Demand Analysis: Estimation of the Quadratic Expenditure System Using Microdata." *Rev. Econ. and Statist.* 66(1984):591–601.
- Barton, A. P. "Family Composition, Prices and Expenditures Patterns." *Econometric Analyses for National Economic Planning: 16th Symposium of the Colston Society*, ed. P. Hart, G. Mills, and J. K. Whitaker. London: Butterworth, 1964.
- Berndt, E. R., and N. E. Savin. "Estimation and Hypothesis Testing in Singular Equation Systems with Autoregressive Disturbances." *Econometrica* 43(1975): 937–57.
- Buchman, S., and S. Gaylann, eds. *CRB Commodity Yearbook*. New York: Commodity Research Bureau, various years.
- Capps, O., J. R. Tedford, and J. Havlicek. "Household Demand for Convenience and Nonconvenience Foods." *Amer. J. Agr. Econ.* 67(1985):862–69.
- Cox, T. L. *A Rotterdam Model Incorporating Advertising Effects: The Case of Canadian Fats and Oils*. Dep. Agr. Econ. Staff Pap. No. 305, University of Wisconsin, May 1989.
- Deaton, A. S., and J. Muellbauer. "An Almost Ideal Demand System." *Amer. Econ. Rev.* 70(1980):312–26.
- Doran, H. E., and J. J. Quilkey. "Harmonic Analysis of

- Seasonal Data: Some Important Properties." *Amer. J. Agr. Econ.* 54(1972):646-51.
- Eales, J. S., and L. J. Unnevehr. "Demand for Beef and Chicken Products: Separability and Structural Change." *Amer. J. Agr. Econ.* 70(1988):521-32.
- Goddard, E. W., and A. K. Amuah. "The Demand for Canadian Fats and Oils: A Case Study of Advertising Effectiveness." *Amer. J. Agr. Econ.* 71(1989):741-49.
- Guthrie, H. A. *Introductory Nutrition*, 6th ed. Santa Clara CA: Times-Mirror Publishing Co., 1986.
- Huang, K. S. *U.S. Demand for Food: A Complete System of Price and Income Effects*. Washington DC: U.S. Department of Agriculture, Econ. Res. Serv. Tech. Bull. No. 1714, Dec. 1985.
- Judge, G. G., W. E. Griffiths, R. C. Hill, and T. C. Lee. *The Theory and Practice of Econometrics*. New York: John Wiley & Sons, 1980.
- Kinnucan, H., and D. Fearon. "Effects of Generic and Brand Advertising of Cheese in New York City with Implications for Allocation of Funds." *N. Cent. J. Agr. Econ.* 8(1986):93-107.
- Lewbel, A. "A Unified Approach to Incorporating Demographic or Other Effects into Demand Systems." *Rev. Econ. Stud.* 52(1985):1-18.
- Moschini, G., and K. D. Meilke. "Modeling the Pattern of Structural Change in U.S. Meat Demand." *Amer. J. Agr. Econ.* 71(1989):253-62.
- Pitt, M. M. "Food Preferences and Nutrition in Rural Bangladesh." *Rev. Econ. and Statist.* 65(1983):105-14.
- Pollak, R. A., and T. J. Wales. "Comparison of the Quadratic Expenditure System and Translog Demand Systems with Alternative Specifications of Demographic Effects." *Econometrica* 48(1980):595-612.
- . "Demographic Variables in Demand Analysis." *Econometrica* 49(1981):1535-51.
- Putnam, J. J. *Food Consumption, Prices, and Expenditures, 1966-1987*. Washington DC: U.S. Department of Agriculture, Econ. Res. Serv. Statist. Bull. No. 773 Jan. 1989.
- Rao, C. R. *Linear Statistical Inference and Its Applications*, New York: John Wiley & Sons, 1973.
- Raper, N., and R. Marston. *Nutrient Content of the U.S. Food Supply*. Washington DC: U.S. Department of Agriculture, Human Nutrition Information Service 1988.
- Ray, R. "A Dynamic Generalization of the Almost Ideal Demand System." *Econ. Letters* 14(1984):235-39.
- . "The Testing and Estimation of Complete Demand Systems on Household Budget Surveys." *Eur. Econ. Rev.* 17(1982):349-69.
- Rossi, N. "Budget Share Demographic Translation and the Aggregate Almost Ideal Demand System." *Eur. Econ. Rev.* 31(1988):1301-18.
- Sahn, D. E. "The Effect of Price and Income Changes on Food-Energy Intake in Sri Lanka." *Econ. Develop. and Cultur. Change* 63(1988):315-40.
- U.S. Bureau of the Census. *Current Population Reports*. Series P-25. Washington DC, various issues.
- . *Statistical Abstract of the U.S.* Washington DC, various issues.
- U.S. Department of Agriculture. *Dairy Situation*. Washington DC, various years.
- . *Fats and Oils Situation*. Washington DC, various years.
- U.S. Department of Agriculture, Agricultural Stabilization and Conservation Service. *CCC Milk Price Support and Related Activities*. Washington DC, various issues 1962-88.
- U.S. Department of Commerce, *Current Industrial Reports: Fats and Oils Production, Consumption, and Stocks*. Washington DC, various years.
- U.S. Department of Health and Human Services, Public Health Service. *The Surgeon General's Report on Nutrition and Health*. DHHS(PHS) Pub. No. 88-50210, Washington DC, 1988.

Books Reviewed

Bowes, Michael D., and John V. Krutilla. *Multiple-Use Management: The Economics of Public Forestlands*. Washington DC: Resources for the Future, 1989, xxiii + 357 pp., \$40.00.

Bowes and Krutilla present theoretical models of multiple-use management for jointly produced services from public forestlands. These models are then used to analyze planning and harvesting decisions for several combinations of jointly produced services from selected U.S. national forests and to assess implications for federal budgeting of jointly produced services from national forests.

Bowes and Krutilla analyze an important policy issue. The 190 million acres in the U.S. National Forest System currently provide a large share of public recreational opportunities and approximately 30% of the domestic timber harvest. In addition to recreation and timber, these forests also provide other services including waterflow control, forage, minerals, and energy. The quantities of each service produced from the national forests are an enduring source of conflict. The management of U.S. national forests continues to be the subject of federal legislation, public planning, and legal actions. A several-fold increase in demand for both recreation and timber from the national forests since World War II has intensified use conflicts.

This book is timely. Historically, economic analysis has played a limited role in U.S. public forestland management. Most public forest decisions were made by foresters trained in the physical sciences. Many of these decisions may be characterized as supply responses to perceived demands. However, in response to federal legislation since 1960, the role of economics in U.S. national forest management has increased. While the opportunities for economic analysis have expanded, standard economic models of the single product, competitive firm are not directly applicable to public forestland management. The large relative size and characteristics of the public forests make analyses of jointly produced services, such as Baumol, Panzar, and Willig, inappropriate. The U.S. Forest Service is not a competitive firm. Competitive markets do not determine timber prices. For many types of recreation, value is difficult to assess, and exclusion is costly. A portion of the public forest is capable of jointly and simultaneously producing two or more services. To remedy the limitations of standard models, the authors explicitly incorporate characteristics of public land management and joint production.

Chapter 1 outlines the history of public forestland management and associated federal legislation in the United States. Chapter 2 describes services from U.S. national forests and trends in demands for those ser-

vices. Of note is the large relative increase in demand for recreation compared to timber from national forests since World War II and the mediocre quality of most national forestland for timber production.

Chapters 3 and 4 develop static and dynamic frameworks, respectively, for multiple-use management. Benefit and cost functions for jointly produced services are specified. In the dynamic model, amenity values are determined by the ages of several adjacent stands.

In chapters 5–10, the multiple-use management framework developed in chapters 1–4 is used to analyze selected planning, harvesting, and budgeting issues. The topics of analysis are the joint production of timber and water in the Upper Colorado River Basin (chap. 5), the effects of timber management on hunting quality in the Black Hills National Forest of South Dakota (chap. 6), the inclusion of both recreation and timber demands in the planning process for the White Mountain National Forest of New Hampshire and Maine (chap. 7), the incompatibility of providing recreation and producing minerals in the White Cloud Peaks of the Challis National Forest (chap. 8), the budget and appropriation process for unpriced resources services on U.S. national forests (chap. 9), and an evaluation of timber sales which do not cover costs on the Shoshone and Mt. Baker-Snoqualmie national forests (chap. 10).

Two important contributions of this book are the quality of the examples and the multiple stand determination of service values. The examples span a range of service combinations, are grounded in the current decision-making framework, and are well analyzed. The authors carefully exploit available data. The consistent use of willingness-to-pay and marginal, rather than total, changes is refreshing. Discussions for two controversial topics, FORPLAN, the linear programming planning model used by the U.S. Forest Service for forest plans, and below-cost timber sales, are balanced.

In most previous forestry models, the age of a single stand determines service values. All stands are harvested at the age that maximizes net present value of land. Here, the ages of more than one stand determine service values. This allows modeling of the joint effects of different aged stands in producing service values. With the consideration of multiple ages, optimal harvesting rules become both more credible and more complex. In contrast to the unique optimal harvest age of single stand models, no general harvest rule for every stand exists. That is, different stands may have different optimal harvest ages. For example, if users prefer a combination of younger and older stands, it may be optimal to harvest some stands more often and to harvest other stands less frequently or not at all.

This book is well written. Topics are well ordered, and transitions are smooth. The graphs are helpful. Throughout, the reader is treated to clear descriptions of standard forestry and resource economics topics. These topics include complements and substitutes, convex costs, the Faustmann Model, the Hartmann Model, maximum sustainable yield (MSY), even-flow timber restrictions, travel costs, and hedonic pricing.

This is an excellent book on a difficult and important policy issue. I expect it to become a supplemental text in many forestry economics, forest policy, and resource economics courses. The book will also interest those who desire an introduction to either the economics of multiple-use or public forestland management.

Richard J. Brazee
Economic Research Service,
U.S. Department of Agriculture

Reference

Baumol, W. J., J. C. Panzar, and R. D. Willig. *Contestable Markets and the Theory of Industry Structure*. New York: Harcourt Brace Jovanovich, 1982.

Braden, John B., and Stephen B. Lovejoy, eds. *Agriculture and Water Quality: International Perspective*. Boulder CO: Lynne Rienner Publishers, 1990, xii + 224 pp., \$25.

Braden and Lovejoy in the introductory comments of this book put the agricultural water quality problem in perspective with the statement: "Pollution from agriculture degrades more streams, more lakes, more ground-water aquifers and more coastal areas of the United States than pollution from any other source. And the United States is not alone" (p. 1). The broad perspective presented in the book is that nations are experiencing increasing difficulties, are unsure of directions to take, and wish to learn from others' experiences. Rather than attempting to deal with all aspects of agricultural water quality issues, the editors, working with a group entitled "Consortium for Agriculture, Resource and Environmental Policy Research," have focused on policy issues. An underlying thought driving the effort is that past policy approaches depending largely on educational programs and appeals for resource stewardship "have been overwhelmed by the capabilities of modern agricultural technologies and by rising public demand for environmental protection" (p. 1). The book is a compilation of ten papers prepared to deal with specific policy issues and cases and is arranged in three parts.

Part one, "Issues and Policy Options," is presented in three chapters. Libby and Boggess start with a chapter entitled, "Agriculture and Water Quality: Where Are We and Why?" which takes on the task of orienting the reader to the broad set of issues and considerations that make up the problem the other

authors deal with in more detail in subsequent chapters. They introduce many physical and technical factors and then focus attention on the policy aspects that will "demand attention by social scientists . . . in years to come" (p. 9). The two chapters following this overview focus on the two policy approaches: incentives and regulation. Segerson first presents the conceptual basis for incentive policies. She provides a thorough treatment of the economic logic underlying *ex ante* and *ex post* incentive mechanisms and includes discussion of a range of specific taxes, subsidies, and liability rules. Anderson et al. "focus on the current and potential use of regulation . . .," and in so doing "survey patterns of current agriculture regulations that have emerged around the world" (p. 63). This chapter also examines the economic theory underlying environmental regulation and compares the efficiency of incentive and regulatory approaches. Both chapters evaluate the strengths and weaknesses of their respective approaches for dealing with surface- and ground-water quality problems resulting from nutrients, pesticides, and soil erosion. All three chapters are thorough, and when taken together provide a useful view of the economic concepts underlying both incentive and regulatory approaches.

Part two, "Policy Applications," consists of five chapters that have as their purpose the introduction of real-world issues and experiences. Kumm describes and analyzes incentive policies, i.e., fertilizer taxes, afforestation subsidies, and intensive extension programs, that have been used in Sweden to reduce nutrient pollution. Denmark is also attempting to reduce nutrient loading from animal wastes and fertilizers and is investigating policies to reduce pesticide use. Dubgaard discusses the proposed tax policies and reports on research by the Danish Institute of Agricultural Economics. In both the Swedish and Danish cases, it was not possible to point to conclusive evidence as to the effectiveness of the various policy approaches.

Reichelderfer, in reviewing U.S. programs, points out that "historically, direct and indirect positive incentives have been the most popular instruments for agricultural (including agroenvironmental) policy implementation" (p. 131). Most federal programs have focused on erosion control, but in recent times more specific surface- and ground-water quality issues are being raised. An important conclusion is "that new incentive policies and programs need to be designed in anticipation of shifting social, economic, and political conditions—not tied to the short-term conditions existing at the time of policy formulation" (p. 142). Helfand and Archibald, in a chapter entitled "California's Proposition 65: A New Regulatory Trend," give insight into these rapidly changing conditions occurring at the state level. Their discussion illustrates the measures voters are willing to take to force policy makers to recognize their environmental concerns.

The last case, by Musgrave, illustrates that often

a formal relationship between a number of jurisdictions is needed to create the basis for dealing with a water quality problem. Australian federal and four state governments had to form a regional compact to deal with the Murray-Darling Basin water problem. The agreement "provides each state with a clear statement of its obligations and rights . . ." but leaves the "states free to pursue their own programs within the resulting framework" (p. 176). Potential for use of economic incentives is seen.

Up to this point the first eight chapters of the book have presented issues and cases from a "multi-national perspective." Only in part three, "Environmental Policies and Agricultural Competitiveness," consisting of two chapters, are international issues and perspectives raised. Runge points up the emerging issue of environmental regulation and protectionism. Recognizing the potential of non-tariff barriers, he calls for GATT negotiations to broaden policy considerations and "prevent protectionism from masquerading as health and safety standards" (p. 195). Young, in the last chapter, further considers trading arrangements, intensity of resource use, and environmental quality, and is the only author to address environmental concerns in less developed countries. He recommends the OECD "Polluter-Pays Principle" as a starting point for international discussions on agricultural trade and environment.

As one would suppose, the book raises many more questions than it answers. But as is often the case, this is the way to make progress. The book provides a good starting point for understanding the major policy issues related to agriculture and water quality and points the way for policy research.

Clyde F. Kiker
University of Florida

Cohn, Theodore H. *The International Politics of Agricultural Trade: Canadian-American Relations in a Global Agricultural Context*. Vancouver, B.C.: The University of British Columbia, 1990, x + 267 pp., \$36.95.

Billed as the first major examination of Canadian-American agricultural trading relations from a political perspective, this volume traces the evolution of U.S.-Canadian trade relations from the 1950s to the 1980s and examines the bilateral and multilateral forces that resulted in the Free Trade Agreement (FTA) that went into effect on 1 January 1989. For American readers, the more interesting aspects may be Cohn's rendition of bilateral trade history with its Canadian perspective; for Canadian readers, the interest may lie more in the forces that shaped the highly controversial FTA. For all readers, the analysis of forces shaping agricultural trade and its growing international importance will be of interest.

Of special interest to classroom teachers will be chapter 3 on "The Organizational Setting" of agricultural trade. After describing the evolution of the

General Agreement on Tariffs and Trade (GATT), Cohn outlines a multiplicity of other international organizations that influence the structure and conduct of agricultural trade. Interwoven into this discussion are the special waivers, such as section 22, that the United States has enjoyed *vis-à-vis* agriculture and the GATT. This is useful information in light of its significance in the recent round of trade negotiations. Cohn also summarizes the agricultural components of the Kennedy and Tokyo Rounds that tried, but largely failed, to make progress on agricultural trade liberalization.

Of interest to the economic academician will be chapter 4, on "International Pricing Arrangements." This chapter outlines the many attempts to establish duopoly and oligopoly pricing in world grain markets. Included are the lesser-known efforts by the United States and Canada to coordinate grain-pricing policies during and after World War II. It was a surprise to this reader, for example, to learn that "The CWB (Canadian Wheat Board) was then (1943) invited to occupy an office in the U.S. Department of Agriculture Building" (p. 67) in Washington. This presence continued until Dwight D. Eisenhower became president in 1953, at which time "the advantages of these close contacts became less evident to his administration" (p. 67).

Cohn then reviews the extended periods of low grain prices that followed, the many unsuccessful attempts to stabilize markets at higher levels and traces out more recent experiences with grain embargoes, export subsidies, and the U.S. plan for world agricultural trade reform. The reactions of political leaders to these efforts makes engaging reading.

Less engaging to American readers will be chapter 5 on "American Surplus Disposal Measures." Here Cohn describes the long series of diplomatic efforts by Canada to alter U.S. behavior regarding grain disposed of outside commercial channels. As Cohn points out, "Canada was a persistent and harsh critic of American tied sales provisions in PL 480 agreements, maintaining that they interfered with the commercial trade of other exporters. The United States, on the other hand, felt that tied sales were an essential component of its food aid and trade policies" (p. 92).

Cohn also describes U.S. efforts to use barter as a means of disposing of surplus commodities, which Canada also opposed, and concludes that "American tied sales and barter were prime illustrations of the vast differences in U.S. and Canadian economic size and capabilities" (p. 110). This theme, that Canada is small in economic size compared to the United States, and by implication that Canada should receive special trade treatment, is echoed in a number of places in the volume. It is not a view that is widely held in Washington agricultural circles.

Of more interest to American readers will be chapter 6, which describes "Canadian and U.S. Agricultural Export Credits." In this case, Canada is portrayed as actively involved in providing trade credit

to other countries. Cohn notes, "the provision of credit normally does not require massive amounts of economic resources. Canada was therefore able to compete with the United States in the export credit sales area, and both countries at various times took the lead in adopting innovative practices" (p. 111). This chapter provides an enlightening description of the role that trade credit has played in expanding world agricultural trade.

The final chapter deals with the issue of "Agricultural Trade Barriers." In a sagacious statement, Cohn observes that "Canadian-American agricultural trade has alternated historically between cycles of protectionism and liberalization" (p. 139). This is a useful reminder, especially for younger readers. It is quite easy to become so immersed in the current minutia of trade policy that one loses sight of how it waxes and wanes over time. We should keep in mind that only in the post-World War II period has trade liberalization been in the ascendancy. Prior to that, protectionism was the order of the day. That is why volumes like this are so valuable. They remind us that we should not take trade liberalization for granted. Over the course of history, free markets and free trade are the exception rather than the norm. Like liberty, the price of free markets and free trade is eternal vigilance. And, like an appreciation of liberty, this starts in the home and classroom.

In this instance, Cohn has provided us with the information to stay vigilant. His volume should be of interest to classrooms and boardrooms that deal with international agricultural trade and especially those with a focus on U.S.-Canadian trade relations. He provides us with an excellent historical review and description of the political and economic forces that have brought the United States and Canada into a closer trading relationship. Based on the reminders in this volume, however, we should not take that relationship for granted. There are forces working to limit it. We can only hope they weaken over time. Perhaps this volume may help achieve that goal.

Leo V. Mayer
U.S. Congress, House of Representatives
Committee on Agriculture

Fuguitt, Glenn V., David L. Brown, and Calvin L. Beale. *Rural and Small Town America*. New York: Russell Sage Foundation, 1989, xxvii + 471 pp., \$55.00.

This volume is one in a continuing series of monographs prepared for the National Committee for Research on the 1980 Census. Written by three noted experts on the demography of rural America, this book provides a thorough and well-documented statistical portrait of rural and small town America circa 1980. Although based primarily upon 1980 and earlier census data, the volume also draws upon national surveys and population estimates for information on the post-1980 period. Aimed at interested public officials

and nonprofessionals, as well as professional social scientists, the book is clearly written and procedures used are carefully described.

The thirteen chapters that constitute the book may be divided roughly into three main categories: (a) population growth and distribution, (b) demographic and socioeconomic characteristics, and (c) economic activities of the population. Chapter 1 describes the organization of the volume and contains a useful discussion of what is meant by rural and small town America. The authors note that their focus is on "(1) the rural population, metropolitan or nonmetropolitan; and (2) the urban nonmetropolitan population" (p. 6). Changes in the census definitions of rural and urban and metropolitan and nonmetropolitan complicate comparisons over time, as does the movement of areas into or out of these classifications. The authors are careful to note and/or adjust for the impact of these changes on their results.

Chapters 2 and 3 describe trends in population distribution and document the pervasive population deconcentration that occurred in the United States during the 1970s. The widely noted "population turnaround," in which nonmetropolitan areas grew more rapidly than metropolitan areas, is analyzed as part of a much larger movement toward population deconcentration that took place at several territorially based levels. The authors note: "This included population decline in the nation's largest cities, a continuing pattern of metropolitan suburbanization, more rapid growth in smaller than larger Metropolitan Statistical Areas, population redistribution away from the densely settled North, and a reversal in growth patterns between metropolitan and nonmetropolitan areas" (p. 100). The reversal of the turnaround in the post-1980 period is also described and major regional differentials in metro-nonmetro growth are documented. What emerges from this analysis is a picture of substantial diversity in growth patterns by major regional and residential categories.

Chapters dealing with demographic and socioeconomic characteristics present a detailed analysis of compositional differences across residential settings. Many of these findings will confirm the reader's expectations with regard to rural-urban and metro-nonmetro populations. However, some findings challenge beliefs about rural America. For example, in spite of higher rural fertility throughout the nation's history, the "oldest population is disproportionately located in nonmetropolitan areas . . ." (p. 137). This is not only the outcome of the selective outmigration of young adults from nonmetro to metro areas, but increasingly since the 1960s, the selective immigration of the elderly to nonmetro locations. Similarly, the stereotypical picture of Hispanic farm workers is belied by the finding that 90% of Hispanics were urban by 1980, making them one of the most urban ethnic groups in the nation. Perhaps less surprising, but no less important, is the continuing significance of residence for household composition and structure. "Rural areas continue to be characterized by more

traditional household structures, in terms of both family living arrangements and life course composition of householders" (p. 183).

The continuing disadvantage of rural and nonmetropolitan residence is ably documented in the chapters dealing with the economic activities and income of the population. Excellent chapters on the labor force and employment, industrial structure, the population associated with farming, and income and poverty reveal the expected patterns. For example, the percentage living in poverty in nonmetropolitan areas in 1986 was 18.1, compared to the metropolitan rate of 12.3. Similar differences favoring metropolitan residents are found for a variety of employment and income measures.

Rural and Small Town America represents an important addition to our understanding of the significance of residence for America's social and economic structure. No reader interested in population distribution or the hierarchy of urban places can afford to ignore this volume. Given the wide range of topics covered and the uniform quality of the coverage, the volume will be of particular interest to agricultural economists interested in rural development, rural sociologists, demographers, and geographers, among others. Two minor cavils distract the reader from what is a valuable contribution to the literature. Chapter 12, on the characteristics of cities, towns, and rural areas, appears curiously inconsistent with the organization of the book. The customary delay in publishing census monographs is likewise discouraging. However, neither of these quibbles seriously diminishes the value of this work which is likely to become the standard reference in the area.

C. Shannon Stokes
Pennsylvania State University

James, W. E., S. Naya, and G. M. Meier. *Asian Development: Economic Success and Policy Lessons*. Madison: University of Wisconsin Press, 1989, 281 pp., \$37.50, \$13.50 paper.

This book reviews and appraises the unprecedented economic development record of countries of Asia. The authors' note states the book "focuses on the extraordinary performance of the Asian newly industrializing countries (NICs) (Hong Kong, Korea, Singapore, and Taiwan), the four resource-based countries of the Association of Southeast Asian Nations (ASEAN-4—Indonesia, Malaysia, The Philippines, and Thailand), and the low-income countries of South Asia (Bangladesh, Burma, India, Nepal, Pakistan and Sri Lanka)" (p. xv).

The book is divided into seven chapters, the first of which provides a comparative overview of each of the countries in the three groupings listed above. The subsequent chapters discuss industrialization and trade policy, savings and development, financial flows and debt, agriculture, human resource development and, finally, policy lessons and prospects.

The major focus of the book is that "a given country's course of development appears not to have been limited by its 'initial conditions,' but to have been determined by the quality of its economic policies" (p. 21). Each of the respective chapters illustrates in detail this concept and provides an empirical data base to support the thesis that what matters most are appropriate policies.

One might surmise this conclusion, given that the study originated from research initiated while the authors were at the Economic Office of the Asian Development Bank (p. xvi). The book is expected to serve a wide audience of economists and noneconomists. It avoids the use of mathematical formulas and keeps technical and theoretical issues at an easily understood level, even for the noneconomist. In this regard, it should be useful to students in business and the social sciences, particularly those concerned with Asian and Pacific areas. Nevertheless, the economic management lessons detailed for the respective countries are broadly applicable in other developing countries.

Among the book's strengths are the numerous, well-organized tables and graphs that increase the book's usefulness to the researcher, student, or government official concerned with understanding and tailoring well-designed investment projects. The comprehensive and current bibliographic references plus the chapter notes are of great value to reacquire agricultural economists with the need to conduct agricultural analysis and policy applications within the broader macroeconomic framework. Chapter 3, which thoroughly explores the role of domestic savings and financial development, provides a classic primer on the need to follow nonrepressive financial policy if a country is to grow rapidly.

One might quarrel with a few minor points, i.e., on page 3 it is stated: "But a quarter century of spectacular economic growth in the Asia-Pacific region has shifted the center of the world economy from the Atlantic to the Pacific." This may be occurring, but as yet the shift is still underway and certainly has yet to be achieved.

On balance, this is an excellent book and valuable for the insights it provides for understanding the differentials of growth among Asian developing countries. It should be considered required reading for all developing country decision makers. While the political and social context of the Asian countries is in several respects obviously different from the situation found in Latin America and Africa, the fiscal, monetary, financial, and trade policies pursued so successfully in the newly industrializing countries represent universal economic principles which, if astutely developed and systematically followed, incorporate the basis for a dynamic evolution to self-sustaining economic growth and wealth.

Alfred Thieme, Jr.
Inter-American Development Bank

Moore, Richard H. *Japanese Agriculture, Patterns of Rural Development*. Boulder: Westview Press, 1990, 337 pp., \$35.00.

The core of this book, and its most valuable contribution, is an intensive case study of Nakada township in the rice-growing Tohoku region of northeast Honshu. Its focus is that of an ethnographer who is sensitive to the persistence of age-old value systems and social structures in a rural community and its (so far successful) effort to survive in a rapidly changing economic environment. The author leads us through the history of the township and one of its villages; its customs and beliefs; the pattern of land ownership and tenancy and the influence of irrigation; the (limited) effects of land reform and countervailing government efforts to promote productive efficiency through land consolidation and leasing; the advent of mechanization, freeing labor for off-farm employment; the survival of both full-time and part-time farms under the protective umbrella of high price supports and subsidies; and the role of the agricultural cooperatives in protecting farmers and managing change.

Throughout the book, microeconomic findings based on the author's field work are intermingled with nation-wide events and policies. A final chapter, added at the urging of the editor, deals with current governmental attempts to accelerate structural changes in order to increase the economic efficiency of the agricultural sector, to reduce support costs, and to open the market to international competition.

It is evident throughout that the author shares the traditional rural values and takes a dim view of any attempt to promote or even facilitate structural change in agriculture. To be sure, he does applaud the growth of off-farm employment, which he rightly credits with helping to preserve the rural communities (although he fears that this escape hatch may be gradually closed by a decline in industrial subcontracting). But he does not accept the economists' prescription that the future of Japanese agriculture lies in larger-scale farming employing fewer farmers and a shift away from rice to livestock products, vegetables, and fruits for which consumer demand is growing and in which Japan is, or can become, internationally competitive.

The author is not comfortable with economic analysis. In any event, he makes clear that his priorities are different. "The focus needs to be redirected from the narrow logic of free trade rules or political convenience to the nutritive value of the Japanese diet, the ecology of food, and its cultural significance" (p. 1). "The central components of any solution to the farm dilemma in Japan must involve the cultural right of the members of the farm household to perpetuate their corporate body through succession, inheritance of land, and ancestor worship" (p. 272). He sees "the main problems encountered by rural society" in "the bride shortage and aging rural population" and in "growing disparities . . . in job security" (pp. 272-73). He marshals environmental and dietary arguments and the cultural symbolism attached to rice in defense of the status quo. On the whole, the book is

more helpful in understanding the persistence of a highly fragmented, rice-based agricultural structure in Japan (including the reluctance of farmers to sell, or even lease land) than in providing a guide to its adaptation to changing economic conditions.

Although the book contains much interesting information, frequent digressions and repetitions make it difficult to read. The presentation could have been improved by more rigorous editing. An overview chapter would have helped to give the nonexpert reader a feel for the broad historical sweep of the evolution of Japanese agriculture before immersing him into specific aspects as seen from the grassroots.

Fred H. Sanderson

*National Center for Food and Agricultural Policy,
Resources for the Future*

Pearce, D. W., and R. K. Turner. *Economics of Natural Resources and the Environment*. Baltimore MD: Johns Hopkins University Press, 1990, 378 pp., \$42.50, \$19.50 paper.

Public interest in environmental matters has enjoyed a recent resurgence. It will not be surprising to find that enrollments in undergraduate courses in environmental and natural resource economics will rise with this resurgence. It is to this audience that Pearce and Turner offer their volume. It is a difficult audience to satisfy given that it has a very diverse range of disciplinary backgrounds. Some potential readers will be already skilled in economic principles and analysis. Others will have backgrounds more in the physical sciences.

Acknowledging this difficulty, Pearce and Turner have attempted to provide a text that strikes a balance. They seek to satisfy the economist with sufficient mainstream analysis, while not deterring the ecologist from pursuing the subject. They seek to offer an intellectual challenge to the economist while giving enough practical relevance to maintain the noneconomist's interest.

The approach used in striking this balance is to follow a distinct pattern of analysis. First, the authors go to some length to begin their various analyses using widely understood statements of environmental issues. They then proceed to "translate" these generalized statements into specific economic formulations of the issues. Once so formulated, relatively standard economic analysis is applied.

The results of the approach will not please all. A good example of problems faced by the approach is the treatment of that currently omnipresent concept, sustainable resource use. The translation offered by Pearce and Turner of this infrequently defined concept is useful in that it helps both economists and noneconomists to understand exactly what is being considered. However the economist's translation in terms of optimal rates of depletion of the natural capital stock will no doubt be disturbing for environmentalists. Similarly, the conclusion reached by Pearce

and Turner that sustainability involves the maintenance of the currently available stock of natural capital will be disturbing to many economists. The authors propose a "sustainability requirement" (p. 225), whereby any reductions to the natural capital stock would need to be compensated for by additions to that stock achieved by other projects within the portfolio of decisions being made. These augmenting projects "would not be required to pass any test relating costs to benefits" (p. 225). While the ardent environmentalist may find this proposition appealing, most economists will see it as a prescription for the creation of a poorer society, less able to afford environmental protection.

A second problem created by the approach relates to the breadth of material that is covered. Words devoted to the balancing of environmental and economic approaches are words not used to discuss other important issues. For instance, scant regard is paid to the office choice theory explanation of the popularity of the strongly interventionist approach to pollution control. A consideration of the importance of the lobbying of special interest groups and the roles bureaucrats and politicians versus the impotency of the individual voter would have added a useful dimension. Similarly, there is little attention given to institutional arrangements that have caused the breakdown of the property rights approach to the resource allocation process.

Another notable omission is a systematic treatment of benefit-cost analysis. As the book progresses, more and more reference is made to the measurement and weighing up of benefits and costs, yet nowhere are the basic framework and fundamental assumptions of benefit-cost analysis established. To the uninitiated this could be confusing.

However, there are also distinct advantages of the approach used. Global pollution problems and policies are often overlooked by texts, but this volume devotes a complete chapter to their consideration. Likewise, the volume extends the analysis to cover centrally planned and developing countries. With such interest in all of these issues, the book's breadth of coverage is welcome.

The authors' style is clear and concise, particularly in the analytical sections of the book. The clarity is aided by the use of diagrams that are uncluttered and simple to follow. The conciseness is aided by the book's structure: chapters tend to be short and address very specific topics, so that ideas and concepts are tightly packaged. The subdivision of chapters shows to full advantage the organization of the underlying argument. These features will enhance the ability of students to assimilate the material covered by the book. Its limited use of mathematics will delight some and frustrate others.

The volume recognizes that noneconomists are not only concerned about the environment but are also interested in finding out what and how economists think about environmental issues. Its analysis is not condescending to noneconomist readers in that it is not overly pedagogical. Yet it has enough footholds

back to environmental basics to make the nonspecialist feel comfortable. For the economist, there is sufficient analytical strength to provide a stimulating challenge. Part of that challenge will be coming to terms with the views of physical scientists and seeking solutions to the environmental problems they pose.

Jeff W. Bennett

University of New South Wales, Australia

Rawski, Thomas G. *Economic Growth in Prewar China*. Berkeley: University of California Press, xxxiii + 448 pp, \$48.00.

This book is required reading for those interested in China's economic development, be they economists, historians, researchers, teachers, or students. Rawski has set out to expose conventional beliefs about China's modern economic history to the harsh light of empirical evidence. The evidence he presents is likely to change, or at least raise serious questions about, those beliefs.

Rawski challenges the view that China's economy was in a state of stagnation during the prewar period. He presents new estimates of growth rates for China's GDP between 1914–18 and 1931–36. These estimates imply per capita GDP growth at rates of about 1% a year, or possibly higher. Such growth exceeds previous estimates by K. C. Yeh and Dwight Perkins of less than 0.5% a year, and it resembles growth achieved by Japan during 1897–31, a period of accelerating growth. Rawski's estimates imply that the economy was not stagnating: real per capita income rose over 20% between 1914–18 and the 1930s.

A wide range of related evidence is used to back up these estimates. Data are most abundant for, and the book focuses on, the modern sector. Separate chapters are devoted to manufacturing, banking and the monetary system, and transport and communications. For each of these sectors, Rawski weaves together data from diverse primary and secondary sources. The picture that emerges is one of dynamism, where modern firms grew and spread regionally and where modern and local institutions frequently achieved a symbiotic relationship. This picture contrasts with the conventional view that the modern sector was limited in influence, did not penetrate beyond a few major urban centers, and grew at the expense of traditional industries.

A chapter on investment examines whether the investment levels were high enough to initiate a long-term growth process. W. Arthur Lewis and Walt Rostow have suggested that "take-off" involves a transition from low investment at or below 5% of gross national product (GNP) to high investment at 10% or more of GNP. Earlier work on China by K. C. Yeh shows gross fixed investment in the 1930s at only 5% of national product. In contrast, Rawski estimates that by that time fixed investment had risen to about 10% of GDP. If correct, these findings contradict beliefs that rates of return on investment were low, that uncertainty and instability prevented investment, and that

for cultural or social reasons China's elites were unwilling to invest.

The credibility of Rawski's arguments rests on the evidence. Information is most abundant and his analysis strongest for the modern sectors. The modern sectors, however, accounted for only 10%–15% of national product. Conclusions about the overall economy, then, require estimates for growth and investment in traditional sectors, especially agriculture. Unfortunately, systematic, nationwide data for output and investment in agriculture do not exist. Rawski therefore uses indirect evidence. Data showing rising cotton cloth consumption and rising wages of unskilled labor from rural areas suggest that incomes and living standards were increasing in rural China. Under certain assumptions, these data imply rising farm output.

Despite the careful discussion of these indirect data, Rawski's conclusions about trends in agriculture, and thus for the economy as a whole, are suggestive rather than definitive. Nevertheless, his study brings important new evidence to bear on the subject. Moreover, if one is uncomfortable with Rawski's estimates for agricultural growth, one must be equally (if not more) uncomfortable with previous estimates. Earlier studies face the same data problems and rely on equally strong assumptions.

I should add that my favorite chapter is the first one, in which Rawski demolishes the view that the prewar Chinese economy was ravaged by militarism and a corrupt, extractive government. Militarism and the government may very well have disrupted the economy in particular regions at particular times. Yet, data show that the numbers of troops, casualties, and military outlays were small relative to the population and economy as a whole. Similarly, data for the 1930s show central government expenditures averaged less than 3% of GDP, and Rawski estimates total government spending (all levels) at between 4.7% and 7.2% of GDP. The government's ability to raise taxes was apparently weak, and so its ability to "extract" resources from the economy was probably limited.

By insisting on systematic evidence as the basis of our views, Rawski has raised the level of debate about China's prewar development. Those who disagree with Rawski's conclusions will either have to refute the data he presents, provide alternative evidence, or challenge his interpretations of the data. Regardless of whether they agree or disagree, readers will find this book a gold mine of primary source references and materials, and a very stimulating work.

Terry Sicular
Harvard University

Reitsma, H. A., and J. M. G. Kleinpenning. *The Third World in Perspective*, 2nd ed. The Netherlands: Van Gorcum, Assen/Maastricht, 1989, xxii + 435 pp., price unknown.

This book was written by two geographers, and it covers the politics, the economics, the sociology, and the legal aspects of development and underdevelop-

ment in addition to geographical aspects. A hint at the breadth of coverage is given by a brief synopsis on the outside back cover of the book. Issues to be described and explained include low productivity, food shortages, poverty, infrastructural deficiencies, underutilization of human and natural resources, population explosion, unemployment, rural-to-urban migration, unbalanced spatial structures, shantytowns, oversized informal sector, inequality, internal colonialism, and corruption. It is designed as an undergraduate textbook but may also serve as supplementary reading for graduate courses. It is rich in illustration—maps, graphs, and photos.

The book is organized into five parts, including twenty chapters and an epilogue. Part one looks at changing perceptions of the third world's underdevelopment beginning with the decade 1945–55 and ending with the decade 1975–85. Part two treats the rural-agricultural scene, while part three covers the urban-industrial scene. Part four's three chapters cover various theories of development and underdevelopment including critiques—especially of the theories of imperialism and dependency. Finally, part five is devoted to a fairly detailed analysis of four country case studies: Ethiopia, India, Cuba, and Taiwan.

This is an excellent book. It is well written and easily comprehended. In their critiques of some theories of development and underdevelopment, the authors make every effort to present a balanced view. They conclude that during the 1960s and 1970s colonialism was overemphasized as a major negative influence. Reitsma and Kleinpenning try to balance this view by giving more attention to the internal factors of the physical and cultural environment operating throughout the centuries. The case studies were selected to provide some comparative testing of a number of the theoretical concepts of development and underdevelopment. In the epilogue, the authors provide their rationale for the countries selected for their case studies, and they present some interesting contrasts with respect to physical and cultural environmental factors, the role of indigenous elites, colonialism, and others.

With respect to colonialism, they conclude as follows: "With the exception of five years of Italian occupation, Ethiopia never was a colonial possession. Today it is one of the world's poorest countries. After fifty years of colonial rule by the Japanese, Taiwan is presently well on its way to becoming a prosperous and developed country. Both Cuba and India were colonial dependencies for a very long period of time, but present living conditions of the 'average' inhabitant are far better in the former than in the latter country" (p. 393).

In chapter 14, "Causes of Underdevelopment: An Overview of Theories" (pp. 211–38), the authors draw heavily on the views of two European geographers, Hans Bobek from Austria and Yves Lacoste from France. Both of these writers have serious reservations concerning the one-sidedness of the anticolonialism interpretation. In fact, say Reitsma and Kleinpenning, "Another point in Lacoste's favor is the

emphasis he places on the harmful role played by colonialism while at the same time bringing out the relative importance of that same factor. By drawing attention to the part played by third world elites, he makes more understandable the lack of progress during the postcolonial period as well as the limited development in third world countries which never were colonies" (p. 235).

The book does have a few flaws in my judgment. On page 33, Reitsma and Kleinpenning speak of the problem of "an excessively large and for the most part unproductive, even parasitical, informal tertiary sector." This is certainly an incomplete view. Much later (p. 164), they, in fact, speak much more positively of this informal sector. Unfortunately, neither De Soto nor Hirschman appears in their extensive list of references.

In several instances they seem to focus on Africa and south-southeast Asia but generalize for the third world as a whole. For example, "In nearly every Third World country, over half the working people are engaged in agriculture and in many at least 75 percent are so employed" (p. 52). This is an inaccurate generalization for Latin America. The sharp dichotomy between large agricultural enterprises producing for the market and subsistence farmers who can barely feed themselves from their production (pp. 96, 102) seems overdrawn for most areas of the less developed world. The Green Revolution is, I believe, underrated (pp. 80-85), and no mention is made of the network of international agricultural research centers. Finally, I fail to understand the need for a questionable comparison which is found on the very first page (p. 3) of chapter 1: "Partly because Europeans are exposed to more information on the low quality of life in the Third World and partly because they live in a more socialistically minded environment, they tend to be more inclined than Americans to sympathize with, and provide economic support to, the poor and hungry in Africa, Asia and Latin America."

Despite several overgeneralizations, this is a good book and should be a useful supplementary reader for courses in development studies.

Peter Dörner
University of Wisconsin

References

- De Soto, Hernando. *The Other Path: The Invisible Revolution in the Third World*. New York: Harper and Row, 1989.
- Hirschman, Albert O. *Getting Ahead Collectively*. New York: Pergamon Press, 1984.

Sicular, Terry, ed. *Food Price Policy in Asia: A Comparative Study*. Ithaca NY: Cornell University Press, 1989, viii + 307 pp., \$49.95, \$16.95 paper.

"The aim of this book is to understand food pricing and through it broader food policy choices in a variety of country settings . . . the chapters have been written to a common outline and share a common perspective . . . food price policy is defined to include both explicit price interventions and macroeconomic, exchange-rate, and trade policies" (p. vii). The six Asian country studies, written by authors with considerable research experience in their respective areas of expertise, are China, Indonesia, Nepal, the Philippines, South Korea, and Thailand. The latter half of the above definition of price policy is critical for the chapter outlines, which, according to Sicular, draw heavily from the ideas of Timmer, Falcon, and Pearson in *Food Policy Analysis*; the effects of these economy-wide policies are dealt with at some length in all of the chapters and constitute the main focus of attention in the Korean and Philippine chapters. The common outline includes a discussion of (a) policy objectives; (b) price policy instruments (output price supports, consumer subsidies, input price subsidies usually for fertilizer, macroeconomic, and trade policies); and (c) policy outcomes and tradeoffs between conflicting policy objectives (urban consumption, producer incomes, price stability, self-sufficiency and food security, government revenues, rapid economic development).

A central question addressed by the authors is whether "distorted" agricultural prices hinder agricultural and overall economic development—how important is "getting prices right?" Kym Anderson (Korea, where agriculture is protected) and Christina David (the Philippines, where agriculture is taxed) take strong positions that past price distortions resulted in serious inefficiencies. Peter Timmer contends that most of the benefits of the agricultural research and investment program in Indonesia would not have been realized without implementation of the more favorable price environment after the mid-1960s that substantially reduced but did not entirely eliminate biases against agriculture. Theodore Panayotou (Thailand) argues that agriculture grew strongly despite substantial taxation of rice exports because of the compensating factors of free irrigation water and free land for expansion; however, the rising costs of bringing new land under cultivation and irrigation will not permit a continuation of past pricing policies with the same favorable outcome.

In Nepal, with its long, open border with India, the question is to some extent irrelevant. Michael Wallace presents the apt analogy of the flea (Nepal) that travels in the elephant's (India) direction, whichever way the flea turns on the elephant's back. Terry Sicular chooses not to criticize directly (or praise) China's policy of taxing food production through various quota arrangements but instead is satisfied to lay out the resulting tradeoffs in policy outcomes.

The brief concluding chapter is disappointing in that Sicular is content merely to summarize the complexities of food price policy formulation—in some sense recapitulating the objectives-instruments-outcomes

framework of Timmer, Falcon, and Pearson with specific examples from the case studies. A comparative analysis of several empirical applications of that framework might have identified common tendencies that would have allowed predictions/rules/hypotheses of government behavior in carefully defined situations. After providing some relevant cross-country background information, a much longer introductory chapter makes a start at this, but the concluding chapter fails to follow up.

The job of identifying these common tendencies would be made easier in my opinion by a more in-depth treatment in the case studies of three areas. First, the development and dissemination of modern varieties of rice allowed policy makers in three countries to compensate rice producers for rice prices which were kept artificially low. Second, relatively little empirical information is provided on consumption effects; it would have been useful to discuss urban-rural tradeoffs in terms of four groups: urban and rural poor and urban and rural wealthy. Third, it is these latter two groups that usually vie for political influence along with the public bureaucracy, the military, and international interests; a short summary of the mechanisms through which the political system in each of these countries mediates these competing concerns would have been helpful.

The strength of the book lies in the wealth of information that is contained in the six country studies, which organize and summarize past research (often the author's own research) on a broad range of country-specific food price policy issues. Each chapter provides a valuable reference for persons interested in food policies in that country. The Indonesian chapter is exemplary in its even treatment of the large number of factors that can bear upon food policy formation, blending theory and empirical evidence into a compelling, understandable story despite its complexity. The Korean chapter is also noteworthy for the persuasive empirical evidence that it presents in measuring the costs of agricultural protection and for its elucidation of why particular policies emerge from the political process.

Howarth E. Bouis

International Food Policy Research Institute

References

- Timmer, C. Peter, Walter P. Falcon, and Scott R. Pearson.
Food Policy Analysis. Baltimore MD: Johns Hopkins University Press, 1983.

Industry Members and Representatives—1990

American Agricultural Economics Association

American Crystal Sugar Co.

101 North Third Street
Moorhead MN 56560-1990
Rep.: David Berg

Cargill, Inc.

P.O. Box 9300
Minneapolis MN 55440
Rep.: Robbin S. Johnson

Central Soya Company, Inc.

P.O. Box 1400
Fort Wayne IN 46801-1400
Rep.: C. Lockwood Marine

Chicago Board of Trade

Education and Marketing Department
141 West Jackson Boulevard
Chicago IL 60604
Rep.: Richard Jelinek

ConAgra, Inc.

Economic Research Department
One Central Park Plaza
Omaha NE 68131
Rep.: Richard L. Gady

Connell Commodities Company

45 Cardinal Drive
Westfield NJ 07092
Rep.: James D. Sullivan

Dairymen, Inc.

10140 Linn Station Road
Louisville KY 40223
Rep.: George H. Jung

Eli Lilly & Company

112 Chesterfield Drive
Noblesville IN 46060
Rep.: Thomas E. Elam

General Mills, Inc.

P.O. Box 1113
Minneapolis MN 55440
Rep.: William J. Vollink

Holt, Miller & Associates

2111 Wilson Boulevard, No. 531
Arlington VA 22201
Rep.: James S. Holt

Kraft USA

Economic Research O & T
2211 Sanders Road
Northbrook IL 60062
Rep.: Marcia E. Glenn

Louis Dreyfus Corporation

Economic Research Department
10 Westport Road, P.O. Box 810
Wilton CT 06897-0810
Rep.: Harlan Burnstein

Oscar Mayer Foods Corporation

P.O. Box 7188
Madison WI 53707
Rep.: Patrick J. Luby

Regi Des Assurances Agricoles

113 St. Georges Street W
Sevisoy, Quebec
Canada G6V 4L2
Rep.: Dir Etudes Economique

The Royal Bank of Canada

Agricultural Banking Services
P.O. Box 923
Winnipeg, Manitoba MB
Canada R3C 2T5
Rep.: John J. Murphy

Area-Yield Crop Insurance Reconsidered

Mario J. Miranda

One of the more promising proposals for reforming the federal crop insurance program calls for both premium rates and indemnities to be based not on the producer's individual yield but rather on the aggregate yield of a surrounding area. Area-yield crop insurance can provide more effective yield-loss coverage than individually tailored insurance, without most of the adverse selection and moral hazard problems that have historically undermined the actuarial performance of the federal crop insurance program.

Key words: crop insurance, optimal hedging, risk and uncertainty.

When the Federal Crop Insurance Act of 1980 was signed into law, policy makers envisioned a crop insurance program that would ultimately operate on a near actuarially sound basis with limited government financial assistance.¹ Between 1980 and 1988, however, government outlays for the federal crop insurance program exceeded 4.2 billion dollars, accounting for over 80% of the total indemnities paid to producers. The loss ratio, indemnities paid to producers divided by premiums collected from producers, averaged 2.05, well in excess of the approximate 0.95 level generally regarded as necessary for break-even insurance operations (U.S. GAO).¹ The poor actuarial performance of the federal crop insurance program and its failure to attract producer participation has led to dissatisfaction with the program, including calls for the repeal of the 1980 act and elimination of federal crop insurance.²

Under the provisions of the 1980 act, crop insurance is marketed primarily through private

insurance agents and brokers. The 1980 act authorized the Federal Crop Insurance Corporation to subsidize producer premium payments and to reimburse participating private insurance companies for their administrative expenditures and part of their underwriting losses. Federal crop insurance, which is available for over fifty crops, covers all natural risks, including unavoidable losses from drought, excessive rain, and storm damage. An agricultural producer can purchase individualized coverage for either 50%, 65%, or 75% of the normal yield, and at one of three different price elections. If the producer's yield falls below the elected coverage level, he receives, per insured acre, an indemnity equal to the yield shortfall times the elected price level.

The failure of the federal crop insurance program to operate on an actuarially sound basis can be attributed to the problems inherent in trying to tailor coverage to individual yield-loss experience. The most serious of these problems, adverse selection, arises because producers are better informed about the distribution of their own yields and thus are better able to assess the actuarial fairness of their premiums than the insurer, who lacks access to reliable individual yield data and other relevant information (Skees and Reed). Producers who recognize that their expected indemnities exceed their premiums are more likely to purchase coverage than those whose premiums are actuarially high. As a result, the insurer's expected indemnity outlays exceed total premium income, and, in the long run, the insurance operation loses money. Efforts by the insurer to avoid these losses by raising premiums only result in a smaller and more adversely selected pool of participants.

Mario J. Miranda is an assistant professor in the Department of Agricultural Economics and Rural Sociology, Ohio State University.

Research was supported by U.S. Department of Agriculture Cooperative Research Agreement 58-3AEK-7-80020.

The author wishes to thank Joseph Glauber, Jerry Skees, and an anonymous reviewer for their valuable suggestions and comments and Jean Buzby for her assistance with the data.

¹ If government premium subsidies are included in the total premium payment, as is common practice, the loss ratio would be 1.57. Excluding the subsidies, as occurs here, provides a better measure of the current program's failure to operate on an actuarially sound basis. The 0.95 loss ratio target assumes that a loading factor of 5% is sufficient to cover administrative expenses.

² The Bush administration's 1990 farm bill proposal calls for repeal of the Federal Crop Insurance Program and replacing it with a standing disaster assistance program.

Other problems associated with individual-yield crop insurance include high administrative costs and moral hazard. Record keeping and other manpower requirements needed to verify individual production histories and to adjust individual yield-loss claims raise insurer expenditures and impose transactions costs on participating producers. Moral hazard occurs when producers, after purchasing insurance, alter their production practices in a manner that increases their chances of collecting an indemnity (Chambers, Nelson and Loehman). In order to combat moral hazard, federal crop insurance requires a deductible of at least 25% of the producer's normal yield. This provision limits the coverage provided by the insurance and reduces its value to the individual producer.

Area-Yield Crop Insurance

The fundamental problems that accompany individual-yield crop insurance have been known since the early days of the federal crop insurance program. Halcrow, in his 1949 evaluation of the effectiveness of the federal crop insurance program's first decade of operation, concluded that individual-yield crop insurance "will work in a satisfactory manner only under a system of conditions so exacting in their specification that they will be found to a rather limited extent in American agriculture" (p. 476).

In his paper, Halcrow promoted an alternative crop insurance scheme in which both indemnities and premiums would be based not on a producer's individual yield but rather on the aggregate yield of a surrounding geographical area. Under a so-called area-yield plan, a participating producer would receive, in any given year, an indemnity equal to the difference, if positive, between the area yield and some predetermined critical yield level. Every participating producer in a given area would receive the same indemnity per insured acre, regardless of his own crop yield, and therefore would pay the same premium rate.³

Area-yield crop insurance offers numerous

advantages over individual-yield crop insurance. Because information regarding the distribution of the area yield is generally available and more reliable than information regarding the distributions of individual yields, insurers could more accurately assess the actuarial fairness of premiums under an area-yield policy, thereby significantly reducing adverse selection problems.⁴ Moreover, because the indemnities would be based on the area yield rather than the producer's yield, a producer could not significantly increase his indemnity by unilaterally altering his production practices. Thus, under an area-yield insurance program, moral hazard essentially would be eliminated. Administrative costs would also be substantially reduced under an area-yield program because claims would not have to be adjusted individually and verification of individual production histories would no longer be required.⁵

In the following sections, a theoretical framework for evaluating the effectiveness and equity of area-yield crop insurance is developed and applied to western Kentucky soybean producers in an empirical illustration. Several questions are addressed: How efficiently, relative to individual-yield crop insurance, does area-yield crop insurance cover individual yield risk? How does the risk reduction effectiveness of area-yield crop insurance vary across producers? How can a producer optimize coverage under an area-yield plan? The article concludes with recommendations on how area-yield crop insurance might be implemented.

Theoretical Analysis

Consider a producer i whose yield \tilde{y}_i is random due to the uncertain effects of weather and other natural phenomena. Suppose the producer operates in an area where the average yield across all farms is \bar{y} . By orthogonally projecting the producer's individual yield \tilde{y}_i onto the area yield \bar{y} , an identity that relates the two is obtained:

$$(\cdot) \quad \tilde{y}_i = \mu_i + \beta_i \cdot (\bar{y} - \mu) + \tilde{\epsilon}_i.$$

³ The program examined in this paper should not be confused with the Federal Crop Insurance Corporation's practice prior to 1981 of writing coverage provisions using area yield data while basing indemnities on individual yield experience. We adopt Halcrow's original usage of the term "area-yield" to refer to a program in which coverage, premiums, and indemnities are all based on area yield experience.

⁴ These include intertemporal adverse selection problems, which arise under the current program when producers make participation decisions based on information that they hold privately at planting time, such as the soil moisture levels on their own land.

⁵ In its recent study, the Commission for the Improvement of the Federal Crop Insurance Program cites many of these advantages and recommends that an area-yield crop insurance pilot program be implemented on an experimental basis in selected areas.

Here,

$$(2) \quad \beta_i = \text{Cov}(\tilde{y}_i, \bar{y}) / \sigma_{\bar{y}}^2$$

$$(3) \quad E\tilde{\epsilon}_i = 0 \quad \text{Var}(\tilde{\epsilon}_i) = \sigma_{\tilde{\epsilon}_i}^2 \quad \text{Cov}(\bar{y}, \tilde{\epsilon}_i) = 0$$

$$(4) \quad E\tilde{y}_i = \mu_i \quad \text{Var}(\tilde{y}_i) = \sigma_{\tilde{y}_i}^2$$

$$(5) \quad E\bar{y} = \mu \quad \text{Var}(\bar{y}) = \sigma_{\bar{y}}^2$$

The coefficient β_i measures the sensitivity of the producer's individual yield to the systemic factors that affect the area yield. Equation (1) decomposes individual yield variation into a systemic component $\beta_i \cdot (\bar{y} - \mu)$ that is perfectly correlated with the area yield and a nonsystemic component $\tilde{\epsilon}_i$ that is uncorrelated with the area yield.

Suppose that the producer is offered area-yield crop insurance in which the indemnity and the premium are both denominated in production units, say, bushels per acre.⁶ The producer purchases coverage at a premium rate of π bushels per acre. If the area yield \bar{y} subsequently falls below a critical yield level y_c , he receives an indemnity \bar{n} , in bushels per insured acre, equal to the shortfall:

$$(6) \quad \bar{n} = \max(y_c - \bar{y}, 0).$$

Assume that the premium π is actuarially fair;⁷ that is, it is equal to the expected indemnity $E\bar{n}$.

With area-yield crop insurance, the producer's net yield equals

$$(7) \quad \bar{y}_i^{\text{net}} = \tilde{y}_i + \bar{n} - \pi,$$

and his yield risk, as measured by the variance of the net yield, equals

$$(8) \quad \text{Var}(\bar{y}_i^{\text{net}}) = \sigma_{\tilde{y}_i}^2 + \sigma_{\bar{n}}^2 + 2 \cdot \text{Cov}(\tilde{y}_i, \bar{n}),$$

where $\sigma_{\bar{n}}^2 = \text{Var}(\bar{n})$ is the variance of the indemnity. By acquiring area-yield insurance, the producer thus reduces his yield risk by an amount

$$(9) \quad \Delta_i = \text{Var}(\tilde{y}_i) - \text{Var}(\bar{y}_i^{\text{net}}) = -\sigma_{\bar{n}}^2 - 2 \cdot \text{Cov}(\tilde{y}_i, \bar{n}).$$

⁶ Although straightforward, generalizing the model to account for price variation would undermine the clarity of the exposition while providing little additional insight into the structure of area-yield insurance.

⁷ Assuming that the premium is actuarially fair allows us to evaluate the insurance coverage solely in terms of its variance reduction, provided we further assume, if only as a first-order approximation, that producers are mean-variance utility maximizers. Meyer has shown that mean-variance decision models are consistent with expected-utility maximization under much weaker behavioral restrictions than had previously been thought.

Assume now that the individual nonsystemic yield component $\tilde{\epsilon}_i$ and the area yield \bar{y} are conditionally independent (a mild assumption given that they are uncorrelated by definition). Then the individual nonsystemic yield component $\tilde{\epsilon}_i$ and the indemnity \bar{n} are uncorrelated, and it follows from (1) that

$$(10) \quad \text{Cov}(\tilde{y}_i, \bar{n}) = \beta_i \cdot \text{Cov}(\bar{y}, \bar{n}).$$

Defining

$$(11) \quad \beta_c = - \frac{\sigma_{\bar{n}}^2}{2 \cdot \text{Cov}(\bar{y}, \bar{n})}$$

and substituting (10) into (9), the risk reduction obtained from area-yield insurance can be rewritten as follows:

$$(12) \quad \Delta_i = \sigma_{\tilde{y}_i}^2 \cdot \left[\frac{\beta_i}{\beta_c} - 1 \right].$$

We refer to β_c as the critical beta. Because the area-yield \bar{y} and the indemnity \bar{n} are negatively correlated, $\beta_c > 0$. Moreover, because the critical beta β_c and the variance of the indemnity $\sigma_{\bar{n}}^2$ are determined by the distribution of the area yield \bar{y} and the critical yield y_c , they are invariant among producers within a given area. It thus follows from (12):

PROPOSITION 1. *For a given critical yield y_c , the risk reduction obtained by producer i from area-yield insurance is completely determined by, and is positively related to, his individual beta, β_i .*

It also follows from (12) that:

PROPOSITION 2. *Area-yield insurance is risk reducing for producer i if and only if $\beta_i > \beta_c$, that is, if and only if his individual beta exceeds the critical beta.*

Thus, producers with high β_i 's can expect significant reduction in yield risk from purchasing area-yield insurance, whereas those with low β_i 's may actually find that area-yield insurance is risk augmenting. A characterization of β_i that is helpful in understanding the significance of the above results is given by

$$(13) \quad \beta_i = \rho_i \cdot \frac{\sigma_{\tilde{y}_i}}{\sigma_{\bar{y}}},$$

where ρ_i is the coefficient of correlation between producer i 's yield \tilde{y}_i and the area yield \bar{y} . As an

immediate consequence of proposition 1 and (13) we have:

PROPOSITION 3. *Ceteris paribus, the more highly correlated a producer's yield is to the area yield, the greater the risk reduction that the producer can obtain from area-yield insurance.*

PROPOSITION 4. *Ceteris paribus, the higher a producer's yield variance, the greater the risk reduction that the producer can obtain from area-yield insurance.*

To ascertain how the risk reduction obtained from area-yield insurance varies across producers within a given area, we must determine how the β_i are distributed within the area and how the critical β_c varies with the critical yield y_c . Although definitive answers to both of these questions can only be obtained empirically, some light can be shed from theoretical considerations.

Consider first the distribution of the β_i 's. If ω_i denotes the proportion of total acreage in the area planted by producer i , then, by definition, $\sum_i \omega_i = 1$ and $\sum_i \omega_i \cdot \bar{y}_i = \bar{y}$, so that

$$(14) \quad \sum_i \omega_i \cdot \text{Cov}(\bar{y}_i, \bar{y}) = \text{Cov}\left(\sum_i \omega_i \cdot \bar{y}_i, \bar{y}\right) \\ = \text{Cov}(\bar{y}, \bar{y}) = \sigma_{\bar{y}}^2.$$

Dividing both sides by $\sigma_{\bar{y}}^2$ and using (2), it follows that

$$(15) \quad \sum_i \omega_i \cdot \beta_i = 1.$$

Thus, the acreage weighted average of the β_i within any area is always one. The dispersion of the β_i and the skewness in the their distribution, on the other hand, may vary among regions and ultimately can only be determined empirically. Intuition suggests that the more homogeneous are the soil and climatic conditions faced by producers in a given area, the more closely the β_i s will cluster around one.

Under mild regularity conditions, the critical beta β_c is an increasing function of the critical yield y_c .⁸ In general, it can be shown that⁹

$$(16) \quad 0 \leq \beta_c \leq 0.5, \text{ and}$$

$$(17) \quad \lim_{y_c \rightarrow 0} \beta_c = 0.0 \quad \text{and} \quad \lim_{y_c \rightarrow \infty} \beta_c = 0.5.$$

⁸ The regularity condition holds for reasonable parameterizations of the lognormal and beta distributions often used to describe yield distributions and has been verified for the empirical distributions of 1974–86 U.S. soybean yields for all U.S. soybean-producing counties.

⁹ Formal proofs of (16) and (17) are available from the author upon request.

It thus follows from proposition 2 that area-yield insurance will be risk reducing for any producer i from whom $\beta_i > 0.5$. Since the average β_i within an area is 1, most producers should find area-yield insurance risk reducing. Area-yield insurance is definitively risk augmenting only if $\beta_i \leq 0$, that is, only if a producer's yield is negatively correlated with the area yield. If $0 < \beta_i \leq 0.5$, area-yield insurance may or may not be risk reducing, depending on the critical yield level y_c ; the higher the critical yield y_c , the more likely that area-yield insurance will be risk reducing.

Until now, we have implicitly assumed that producers cover exactly 100% of their acreage when they purchase crop insurance. Suppose now that producer i is free to elect a coverage level ϕ_i that may be more or less than 100%. At this coverage level, producer i 's net yield is

$$(18) \quad \bar{y}_i^{\text{net}} = \bar{y}_i + \phi_i \cdot \bar{n} - \phi_i \cdot \pi$$

and the risk reduction obtained from area-yield insurance is

$$(19) \quad \Delta_i = \text{Var}(\bar{y}_i) - \text{Var}(\bar{y}_i^{\text{net}}) \\ = -\phi_i^2 \cdot \sigma_{\bar{n}}^2 - 2 \cdot \phi_i \cdot \text{Cov}(\bar{y}_i, \bar{n}).$$

Substituting (10) into (19), the risk reduction can be rewritten more conveniently

$$(20) \quad \Delta_i = \sigma_{\bar{n}}^2 \cdot \left[\frac{\beta_i}{\beta_c} \cdot \phi_i - \phi_i^2 \right].$$

Maximizing this expression with respect to the coverage level, it follows that:

PROPOSITION 5. *If the coverage level is optional under an area-yield plan, then producer i minimizes his yield risk by selecting a coverage level*

$$(21) \quad \phi_i^* = \frac{\beta_i}{2 \cdot \beta_c}.$$

Thus, if the producer is free to choose his coverage level, save that it be positive, he can reduce his yield risk using area-yield insurance if and only if his beta is positive, that is, if and only if his yield is positively correlated with the area yield. In this case, since the critical beta β_c rises with the critical yield y_c , the producer's optimal coverage level will fall with the critical yield. Since the critical beta β_c is bounded above by one-half, the producer's optimal coverage level will approach but will never drop below β_i . And thus, since the average β_i is 1, one can expect that coverage in excess of 100% will be optimal for a significant portion of, if not most, producers.

Substituting (21) into (20) and solving gives the following result:

PROPOSITION 6. *If the coverage level is optional, then the maximum risk reduction that producer i can obtain from area-yield insurance is*

$$(22) \quad \Delta_i^* = \rho^2 \cdot \beta_i^2 \cdot \sigma_{\bar{y}}^2,$$

where ρ is the correlation coefficient between the indemnity \bar{n} and the area yield \bar{y} .

Because the area yield \bar{y} and the individual nonsystemic yield component $\bar{\epsilon}_i$ are uncorrelated, it follows from (1) that

$$(23) \quad \sigma_{\bar{y}_i}^2 = \beta_i^2 \cdot \sigma_{\bar{y}}^2 + \sigma_{\bar{\epsilon}_i}^2.$$

That is, yield risk without crop insurance can be decomposed into a systemic component $\beta_i^2 \cdot \sigma_{\bar{y}}^2$ and a nonsystemic component $\sigma_{\bar{\epsilon}_i}^2$. Since $0 \leq \rho^2 \leq 1$, proposition 6 implies that area-yield crop insurance, in effect, eliminates a portion of systemic yield risk faced by the producer but none of the nonsystemic yield risk. Since ρ^2 is invariant across producers, it follows that:

PROPOSITION 7. *If the coverage level is optional, then the maximum risk reduction that can be obtained from area-yield insurance, as a proportion of systemic yield risk, is the same for every producer.*

Empirical Application

We now illustrate how area-yield crop insurance might perform in practice using individual farm-level yield data for 102 western Kentucky soybean producers.¹⁰ The producers are assumed to comprise the entire population of the "area" in which area-yield crop insurance is offered. All performance estimates are derived directly from the empirical yield distributions; a parametric distribution is not fitted to the data.

Table 1 shows how the critical beta for the 102-farm area varies with the level of the critical yield. Critical yields are expressed both in bushels per acre and as a percentage of the normal or expected area-wide yield of 30.7 bushels per acre. As seen in table 1, the critical beta rises as the critical yield is increased. For sufficiently low critical yields, the critical beta

Table 1. Critical Beta, Per-Acre Premium, and Average Optimal Coverage Under an Area Yield Plan, Selected Critical Yields

Critical Yield		Critical Beta	Premium (bu./acre)	Average Optimal Coverage (%)
Percent of Normal	Bushels per Acre			
60	18.4	0.00	0.00	
65	20.0	0.04	0.06	1329
70	21.5	0.10	0.17	506
75	23.1	0.16	0.27	312
80	24.6	0.22	0.37	227
85	26.1	0.25	0.60	198
90	27.7	0.28	0.96	178
95	29.2	0.31	1.47	160
100	30.7	0.36	2.04	138
105	32.3	0.41	2.77	123
110	33.8	0.44	3.70	112
115	35.4	0.48	4.80	104
120	36.9	0.50	6.18	101
125	38.4	0.50	7.69	100

achieves its theoretical minimum of zero and, for sufficiently high critical yields, achieves its theoretical maximum of one-half.

Table 1 also shows that the actuarially fair premium under a full coverage area-yield plan (or, equivalently, the expected per-acre indemnity) rises with the critical yield level. For sufficiently low critical yields, area-yield insurance is completely ineffective and the fair premium is zero; for sufficiently high critical yields, a one bushel increase in critical yield simply raises the expected indemnity, and therefore the fair premium, by the same amount.

Table 1 also shows how the optimal level of coverage under an optional area-yield plan varies with the critical yield level. Because the optimal coverage level varies among producers, only the average coverage level across producers is reported. For example, given a critical yield equal to 90% of normal, producers, on average, minimize their yield risk by purchasing coverage for 178% of their acreage. That is, on average, producers minimize their yield risk by purchasing insurance for 78% more acreage than they actually plant. If the critical yield is set too low, area-yield insurance will be ineffective and the optimal coverage level will be zero. If the critical yield is set sufficiently high, the average optimal coverage level will equal 100%. In intermediate cases, the average optimal coverage level exceeds 100% and falls as the critical yield rises.

Figure 1 shows the distribution of individual betas for the 102 producers comprising the area

¹⁰ The yield data, which were provided by Jerry Skees of the University of Kentucky, cover the period 1974–88 and were adjusted for secular trends to reflect 1988 production levels.

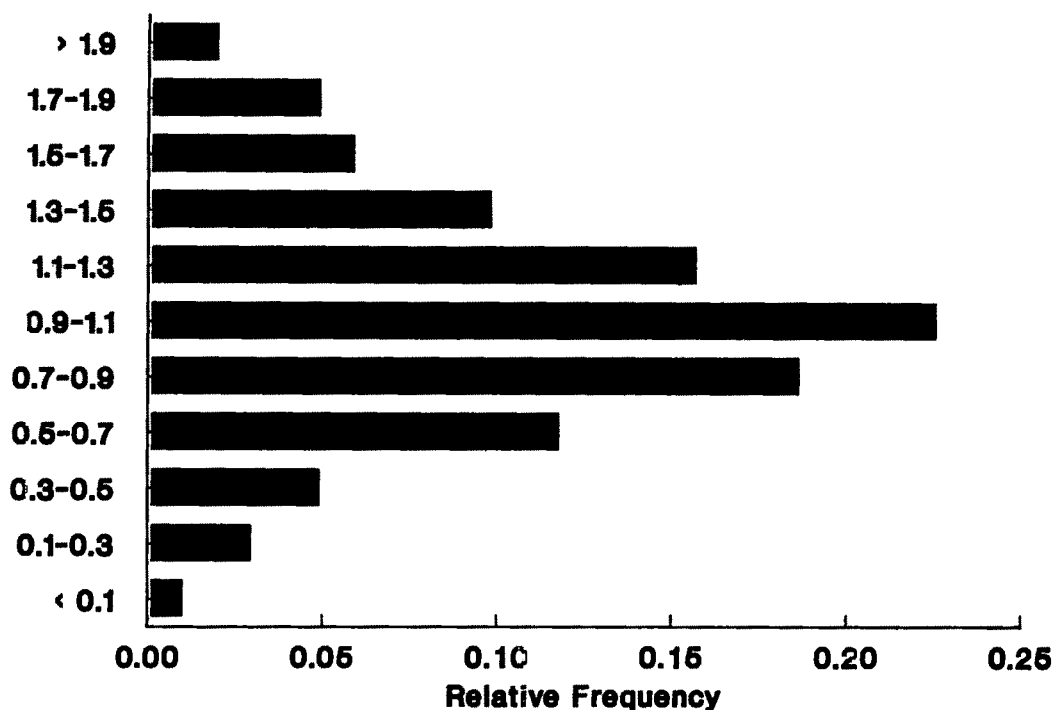
Beta Range

Figure 1. Frequency distribution of individual producer betas

population. The distribution of the betas possesses a regular, bell shape which is centered on one and exhibits no discernable skewness. Of the 102 producers in the sample, 5 have betas falling below 0.28, the critical betas for a full coverage area-yield plan with a critical yield equal to 90% of normal; for those individuals, a full coverage area-yield plan would offer no protection against yield risk. Of the 5 producers, however, 4 have positive betas, indicating that they would obtain some yield risk reduction from an area-yield plan if they could select their levels of coverage optimally.¹¹

We now turn to a comparison of specific versions of an individual-yield plan (IYP) and full coverage and optimal coverage area-yield plans (AYP). Under the IYP, each producer is assumed to choose a yield guarantee equal to 75% of his normal yield, the highest yield election

level available under the current Federal Crop Insurance Program. Under the IYP, whenever a producer's own yield falls below 75% of his normal yield, he receives an indemnity equal to the shortfall.

Under the full coverage AYP, as under the IYP, each producer purchases coverage for exactly 100% of his acreage. The critical yield under the full coverage AYP is set at 88.5% of the normal area yield, indicating that whenever the area yield falls below 88.5% of normal, each producer, regardless of his own yield, receives an indemnity equal to the shortfall in the area yield. A critical yield of 88.5% assures, though only for the present empirical application, that the fair premium paid by each producer under a full coverage AYP is equal to the average fair premium paid by producers under the 75% IYP.¹²

Under the optimal coverage AYP, producers choose their level of coverage and are assumed

¹¹ The 102 farms in the sample are spread over a twenty-county area. Participants in an AYP implemented on a county level would likely be more homogenous with respect to soil and climatic conditions and thus should exhibit less dispersion in their individual betas. One would therefore expect an AYP to perform better than is indicated in this empirical application.

¹² Since individual yields tend to be more variable than the aggregate area yield, an individual yield plan with a given yield guarantee will pay an indemnity more often and thus will require a higher premium than an area-yield plan whose critical yield is the same.

to do so in a manner that minimizes their yield risk. The critical yield under the optimal coverage AYP is assumed to equal 95% of the normal area yield, indicating that whenever the area yield falls below 95% of normal, each producer, regardless of his own yield, receives an indemnity equal to the shortfall in the area yield times his elected level of coverage.

Table 2 shows the actuarially fair premium rates per planted acre under the IYP and the two AYPs for selected producers.¹³ Under the IYP, fair premiums are based on individual yield experience and thus vary among producers. Under a full coverage AYP, the premium is based on the area yield experience and thus is the same for all producers. Under both the IYP and full

coverage AYP, the average premium paid by producers is 0.83 bushels per planted acre, about 2.7% of the normal area yield. Under the optional coverage AYP, the optimal level of coverage per planted acre, and thus the fair premium, varies among producers. The average optimal coverage level across producers is 160%; the additional coverage raises the average premium paid by producers to 2.35 bushels per planted acre.

Table 3 gives the percentage reduction in individual yield risk for selected producers under the three alternative crop insurance plans. As seen in the table 3, risk reduction is greater, on average, under the optimal coverage AYP than under either the IYP or the full coverage AYP. Under the optimal coverage AYP, the average risk reduction is about 39.1%. In contrast, the IYP reduces risk by 30.8% on average and the full coverage AYP by 22.4%.

The relative performance of the three pro-

¹³ Because of adverse selection and heavy government subsidization, most producers under the current crop insurance program pay premiums that are significantly less than their expected indemnity.

Table 2. Crop Insurance Premium Under an Individual-Yield Plan, Full Coverage Area-Yield Plan, and Optimal Coverage Area-Yield Plan, Selected Producers

Number	Beta	Normal Yield ^a	Premium in Bushels per Planted Acre		
			Individual Plan	Full Coverage Area Plan	Optimal Coverage Area Plan
1	2.03	22.5	2.21	0.83	4.77
5	1.84	26.5	2.49	0.83	4.33
9	1.67	29.6	1.82	0.83	3.92
13	1.51	30.1	1.53	0.83	3.55
17	1.41	25.8	1.60	0.83	3.30
21	1.32	26.2	1.55	0.83	3.11
25	1.28	32.7	0.78	0.83	2.99
29	1.23	25.9	1.34	0.83	2.88
33	1.18	27.6	1.21	0.83	2.76
37	1.15	20.9	1.27	0.83	2.69
41	1.08	27.8	0.70	0.83	2.53
45	1.06	27.3	1.52	0.83	2.48
49	1.00	32.3	1.09	0.83	2.36
53	0.98	34.8	0.96	0.83	2.30
57	0.95	27.1	0.42	0.83	2.24
61	0.91	33.0	0.33	0.83	2.13
65	0.87	28.9	0.97	0.83	2.03
69	0.85	27.2	0.51	0.83	1.99
73	0.83	28.7	0.22	0.83	1.94
77	0.79	33.2	0.39	0.83	1.86
81	0.71	23.2	0.76	0.83	1.66
85	0.64	28.3	0.43	0.83	1.50
89	0.56	35.2	0.03	0.83	1.31
93	0.53	18.7	0.98	0.83	1.25
97	0.33	40.4	1.57	0.83	0.78
101	0.27	42.5	0.80	0.83	0.63
Avg.	1.00	30.7	0.83	0.83	2.35

Note: "Selected producers" indicates every fourth producer in order of descending beta value.

^a Average detrended yield, 1974-88, bushels per are, 1988 equivalent.

Table 3. Yield Variance Reduction Under an Individual-Yield Plan, Full Coverage Area-Yield Plan, and Optimal Coverage Area-Yield Plan, Selected Producers

Number	Beta	Yield Variance ^a	Percent Variance Reduction		
			Individual Plan	Full Coverage Area Plan	Optimal Coverage Area Plan
1	2.03	136.4	37.9	20.5	57.3
5	1.84	130.6	50.4	19.2	49.3
9	1.67	112.8	49.4	19.7	46.7
13	1.51	82.5	45.7	23.9	52.5
17	1.41	97.2	35.0	18.5	38.4
21	1.32	74.1	45.0	22.6	44.8
25	1.28	45.6	39.3	35.0	67.5
29	1.23	63.4	44.0	23.9	44.9
33	1.18	56.9	45.0	25.2	45.8
37	1.15	58.1	43.6	23.9	42.6
41	1.08	54.7	27.6	23.4	40.1
45	1.06	94.1	38.9	13.2	22.4
49	1.00	82.5	43.8	14.1	23.1
53	0.98	64.9	37.8	17.3	28.0
57	0.95	34.1	20.6	31.7	50.4
61	0.91	29.8	24.2	34.0	52.6
65	0.87	55.2	40.3	17.1	25.7
69	0.85	29.2	35.6	31.3	46.6
73	0.83	34.1	11.6	25.9	38.0
77	0.79	25.9	34.3	31.9	45.8
81	0.71	57.9	24.0	11.9	16.4
85	0.64	31.4	24.9	18.4	24.5
89	0.56	40.7	1.2	11.1	14.5
93	0.53	42.8	44.3	9.6	12.5
97	0.33	136.8	34.8	0.7	1.5
101	0.27	65.7	41.3	-0.2	2.1
Average	1.00	56.5	30.8	22.4	39.1

Note: See note, table 1.

^a Variance measured in bushels per acre squared.

grams varies among producers. Table 3 confirms that producers with the highest betas tend to enjoy the greatest relative risk reduction under the optimal AYP, whereas producers with the highest yield variances tend to enjoy the greatest relative risk reduction under the IYP. The correlation coefficient between the risk reduction obtained under the IYP and that obtained under the optimal AYP is only 0.06, indicating little relation between the risk reduction benefits from one program as compared to the other. Both small and large producers will tend to prefer the optimal AYP to the IYP, but the latter appear to enjoy the greatest benefit from moving to the optimal AYP from the IYP. For the largest twenty-three producers, who account for half the acreage planted, the optimal AYP offers an average risk reduction of 40.1% and the IYP an average of 28.8%. For the smallest seventy-eight producers, the optimal AYP offers an average risk reduction of 37.3% and the IYP an average of 32.7%.

Additional Considerations

Strictly speaking, an AYP is not a true insurance program since payments to producers are not based on their own specific yield losses. Instead, an AYP is more accurately described as a hedging instrument. Specifically, an AYP is like a put option in which the critical yield plays the role of the strike price. More precisely, because each individual's yield is a constituent of the area yield, an AYP offers a hedge against individual yield-loss in the same way that an option on a stock futures index offers a hedge against the price risk of holding one of the stocks in the index.¹⁴

While characterizing the AYP as a hedging

¹⁴ With slight modification, many of the results in this paper generalize to crop insurance schemes in which indemnities are based on weather variables such as rainfall or temperature. Such schemes have been proposed in the past (Sanderson, Lee) but have failed to gain acceptance among policy makers.

instrument helps put the form and function of the AYP into perspective, it also points to some difficulties that may arise in having the program accepted by producers. Given that producers are notorious for shunning options and futures markets as means of hedging price risk, it is likely that an AYP would not automatically enjoy widespread acceptance among producers as a means of hedging yield risk. Ultimately, to achieve high rates of participation, AYPs may have to be subsidized or made a cross-compliance provision of government commodity stabilization programs.

Should producers be free to select their coverage levels under an AYP, and how high should the critical yield level be set? These two important questions regarding the implementation of an AYP are closely related. Under an IYP, a deductible, or equivalently, a yield guarantee well below the producer's normal yield, is necessary to guard against the moral hazard; for the same reason, coverage levels exceeding 100% of acreage planted cannot be permitted. Under an AYP, however, moral hazard is essentially eliminated; only through illegal and widespread collusive action by producers could the area yield be significantly reduced.¹⁵ Accordingly, optional coverage levels and high critical yields are feasible under an AYP. Since, as shown above, producers can reduce yield risk by choosing coverage levels in excess of 100%, it seems sensible that coverage levels be optional under an AYP. In addition, because the risk reduction under an optimal AYP rises for all individuals as the critical yield is increased, a high critical yield level is indicated.

The higher premiums that producers would pay under an AYP for increased coverage and a higher critical yield need not be a major concern to producers: if the premium is actuarially fair, then producers can expect to recover the higher premium through higher indemnities while enjoying the benefits of reduced yield risk. There are, however, some potential problems that the government may face if critical yield levels are high and optional coverage is permitted under an AYP. First, coverage levels in excess of 100% and critical yields in excess of the normal yield would be difficult to rationalize politically, particularly if the AYP is promoted as an insurance program. Second, higher critical yields and cov-

erage levels would raise the level and variability of total indemnity outlays. Thus, whether the government acts as a direct insurer or reinsurer, substantial yield risk could ultimately be transferred to the government, raising the variability of federal budgetary outlays.

Risk reduction under an optimal AYP varies among producers, raising questions about the equity of such a program. Recall, however, that an optimal coverage AYP would reduce systemic yield risk in the same proportion for all producers. If areas are defined so as to be homogenous with regard to soil and climatic conditions, then nonsystemic yield risk would be attributable almost exclusively to producer-specific factors such as choice of production practice. It is arguable, therefore, that the proper goal for a government crop insurance program should be the reduction of systemic risk, not total risk; otherwise, the program would promote a misallocation of societal resources by encouraging risky production. Thus, if areas are properly defined, an optimal coverage AYP will be equitable in a socially meaningful way.

Because of the abundance of reliable historical yield data and the pre-existence of administrative structures, the most practical definition of area under an AYP will likely be the county. If the AYP is administered on a county basis, the maximum efficacy and equity of the AYP will be achieved in counties that are homogenous with regard to soil and climate. In some regions, however, area boundaries that cut across established county lines may be needed for the AYP to function well. Regardless of how areas are determined, however, it is conceivable that some producers will find the yield of an adjacent area more representative of their own individual yields. In such cases, the producers would improve their coverage by participating in the AYP of the adjacent area and thus should not be discouraged from doing so.

Concluding Comments

An individual producer's total yield risk can be decomposed into a systemic component that is explained by factors affecting all producers in his area and a nonsystemic residual component. Individual-yield crop insurance, such as the one currently underwritten by the federal government, covers total individual yield risk but is limited in its effectiveness by the large deductibles that must be imposed in order to combat moral hazard. Area-yield crop insurance would

¹⁵ A collusive arrangement would be difficult to sustain since a noncolluding producer could maximize his individual yield and still obtain the benefits of the higher indemnity brought about by the actions of the colluding producers.

cover only systemic individual yield risk but would also be free of moral hazard and thus would not require large deductibles or limits on coverage levels. For most producers, the improved coverage of systemic yield risk obtained through lower deductibles and higher coverage under an area-yield plan would outweigh the nonsystemic yield risk protection provided by an individual-yield plan. That is, for most producers, area-yield insurance would provide better overall yield risk protection than individual-yield insurance.

Because information pertaining to the distribution of an area yield is not privately held and is generally available, the asymmetric information problems that have given rise to adverse selection under the current federal crop insurance program would be significantly reduced under an area-yield program. The reduction of adverse selection and the virtual elimination of moral hazard would significantly improve the actuarial performance of the federal crop insurance program. Moreover, because verification of individual production histories and adjustment of individual yield-loss claims would not be necessary under an area-yield program, an area-yield crop insurance program would be less expensive to administer.

The evidence presented in this paper suggests that area-yield crop insurance should receive serious consideration as an alternative to the current crop insurance program. At the very least, area-yield crop insurance should be examined on an experimental basis with long-term pilot programs established for regions where tradi-

tional crop insurance has historically performed poorly.

[Received February 1990; final revision received June 1990.]

References

- Chambers, Robert G. "Insurability and Moral Hazard in Agricultural Insurance Markets." *Amer. J. Agr. Econ.* 71(1989):604-16.
- Commission for the Improvement of the Federal Crop Insurance Program. *Recommendations and Findings to Improve the Federal Crop Insurance Program*. Washington DC, July 1989.
- Halcrow, Harold G. "Actuarial Structures for Crop Insurance." *J. Farm Econ.* 21(1949):418-43.
- Lee, Ivan M. "Temperature Insurance—An Alternative to Frost Insurance in Citrus." *J. Farm Econ.* 35(1953):15-28.
- Meyer, Jack. "Two-Moment Decision Models and Expected Utility Maximization." *Amer. Econ. Rev.* 77(1987):421-30.
- Nelson, Carl H., and Edna T. Loehman. "Further Toward a Theory of Agricultural Insurance." *Amer. J. Agr. Econ.* 69(1987):523-31.
- Sanderson, Fred H. "A Specific-Risk Scheme for Wheat Crop Insurance." *J. Farm Econ.* 25(1943):759-76.
- Skees, Jerry R., and Michael R. Reed. "Rate-Making and Farm-Level Crop Insurance: Implications for Adverse Selection." *Amer. J. Agr. Econ.* 68(1986):653-59.
- U.S. Department of Agriculture, Office of Publishing and Visual Communication. *1990 Farm Bill: Proposal of the Administration*. Washington DC, Feb. 1990.
- U.S. General Accounting Office. "Disaster Assistance: Crop Insurance Can Provide Assistance More Effectively Than Other Programs." Report to the Chairman, Committee on Agriculture, U.S. House of Representatives, Washington DC GAO/RCED-89-211, Sep 1989.

Scientific Principle and Practice in Agricultural Economics: An Historical Review

Harold F. Breimyer

Ever since their field of learning took its own identity, agricultural economists have seen themselves as dedicated to the scientific principle. In early years that commitment was trusted as a guard against human fallibility, and the more so as the first investigations relied heavily on statistical data and analyses—supposedly bias-free. Thus originated a lasting emphasis on the quantitative. A need for underlying economic principles, recognized at once, was developed in depth beginning in the 1950s. Research taxonomies, outlined in the 1920s, devolved later into an array of research methods. Scientific practice is viewed also as how scientists “behave,” and writings about agricultural economists’ record are couched in terms of values, objectivity, ethics. Despite just claims to systematic rigor, science responds also to unsystematic elements, among them imagination and hunch. Most engaging in literature are never-ending exchanges about the merits of mathematics and model building.

Key words: research methods, scientific principle.

Agricultural economists are diverse and can be disputatious, but they have one common bond. They see themselves as functioning in the realm of science and dedicated to the scientific principle.

It is, therefore, something of an anomaly that in the annals of their discipline, inquiries into the meaning of that principle and its significance to their field are reported only episodically. It is as though, having embraced the doctrine of science, practitioners in agricultural economics have felt little need to meditate on how their endeavors fit into the schema of learning known as science, or even to ponder what that schema means. Nor do they anguish over what constraints are thereby placed on them.

With a slight literary license the history of agricultural economics can be said to be a sequence of counting, correlation, and econometrics. The early years of collection of data were followed by a decade of fascination with correlation and then by a long dedication to econometrics with all its model building and intricate statistical analysis. There may be whimsy in the lexicography of the three terms, which are of

Middle English, Latin, and Greek origins, respectively. Is the profession reverting to its roots?

The earliest professional writing in agricultural economics had the simplest of themes. It was to “get the facts,” which in most instances were numerical. That focus led to a major emphasis on data gathering, that proved also to be precedent setting as it established a lasting predilection for quantitative analytical techniques.

In 1963 Don Paarlberg scolded agricultural economists for their bent to quantify everything and omit the nonquantifiable (p. 1390). But quantitative studies continued to crowd non-quantitative ones out of the pages of the *American Journal of Agricultural Economics*, particularly the four quarterly issues.

From the early emphasis on numbers, an inference can be drawn that has never vanished. Subtle and not citable as such in literature, it is a faith that the practices of the scientific world can be so impersonal and trustworthy as virtually to ensure against the perils of human fallibility.

Apparently, during the profession’s first years numbers were seen as pure and their message easily read. Statistical data were regarded as self-identifying as to their relevance and the interpretations to be drawn from them. The literature of perhaps the first quarter century reveals little

Harold F. Breimyer is Professor Emeritus of Agricultural Economics, University of Missouri.

introspective thinking as to what premises or theories necessarily underlie the design of investigatory research or bear on the analytical technique to be applied. It seems incredible that in the 1021 pages of the Taylors' grand *Story of Agricultural Economics* (Henry and Anne Dewees), comments on research method are confined to the briefest of notes on five reports published by the Social Science Research Council. The fragments total scarcely more than a single page (pp. 161, 894, 963–64, 990, 1019–20).

Analytically, during the 1920s the running of averages, frequency distributions, and simple cross-tabulations gave way as the mathematical technique of simple and multiple correlation took the profession by storm. Mordecai Ezekiel's first writing in the *Journal of Farm Economics* appeared in 1923 (pp. 198–213). The *Journal* of the next year carries an article in which Henry Schultz demonstrates the applicability of multiple correlation to a topic that has been copybook for aspiring agricultural economists from that day to this, namely, the demand for beef (pp. 254–78).

The faith that numbers will tell their own story without risk of economists' fallibility was extended to the new technique. It was a shock when Elmer Working pointed out, in a now-classic 1927 article published in the *Quarterly Journal of Economics*, that price-quantity pairings do not even reveal whether a supply or a demand relationship is being described—the identification problem.

Periodically since Working's warning, someone has reminded that there is an element of treachery to the coefficients that correlations produce so readily. In the 1941 and later editions of his price analysis text, Geoffrey Shepherd inveighs against failing to distinguish between correlation and causation (1941, pp. 137–38). A. N. Halter rings the same bell and alleges that economists generally are unclear about causation (1958, p. 1872). In his textbook Shepherd adds a line, almost of apostasy, that “all a correlation coefficient shows is that if the figure is high the relationship is unlikely to be due to chance” (1941, p. 138). As recently as 1988 the correlation-causation issue resurfaced. Thurman and Fisher draw on barnyard examples of the chicken and the egg to treat, half-facetiously, weighty discourses on (alleged) causality in the macroeconomic world (p. 237).

By the time econometric techniques appeared on the scene, the intellectual climate had changed. To be sure, some of the new enthusiasts with

antecedents from early years were quick to pin faith in the miracle-producing power of the new system. But by that time agricultural economists as a genre had become more sophisticated. What followed was a steady flow of exchanges that sometimes dipped into inelegant coinages such as GI–GO (garbage in–garbage out) but more often examined perceptively how numbers-handling or -manipulation is constrained by the precepts of science and the scientific method. A vignette from writings about the econometric era is presented as an epilogue to this review.

In pages that follow will be sketched the principal currents of thought as to the scientific principle in agricultural economics, dating from the profession's earliest years. Sources are primarily the writings in this journal and its predecessors, together with selected other publications.

The Meaning of Science, Place of Theory, and Taxonomy of Research Techniques

The first page of the first issue of the *Journal of Farm Economics* carries an editorial promise that the new publication will “aim to be a seeker for and an expounder of the scientific facts. . . .” In the third issue, published the same year (1919), Richard T. Ely writes of the need for “an agricultural society whose aim is scientific preparation for sound agricultural legislation” (p. 109). Three years later W. J. Spillman, one of the early seminal thinkers, observes that “a science of farm management [lies] hidden in the field” (p. 99). Thus Spillman and Ely, between them, were concerned for science both in managing farms and enacting laws.

But neither they nor their contemporaries addressed the abstract meaning of science. Before long a number of journal writers began to nibble at the edge of the meaning of pure science. E. W. Allen must be credited with being the first contributor to the *Journal of Farm Economics* to take up the cudgel of the meaning of science to agricultural economics. “The aim of science is to discover order, something which occurs with regularity, represents a relationship, or may be expressed in a law. . . . Science assumes order and its task is to discover it” (p. 18).

In 1956 Shepherd dipped briefly into the meaning of science and its bearing on research (pp. 8–14). It remained for Halter in 1958, however, to initiate a roundtable-by-rejoinder (a common pattern) on the topic. Halter found it necessary to begin with the definition of three

words, the first of them science, "meaning that type of mental activity" which satisfies the conditions of "orderly or systematic thinking" and "a definite subject-matter" (p. 1871). Two years later Halter wrote again, incurring Back's rejection of what he sees as Halter's "purely mechanical theory of human conduct" (1960, p. 1474); whereupon Conklin contends that agricultural economists have overly "abstracted from flesh and blood men" and have done so as they "idolized Newtonian mechanical concepts as the acme of science. . . ." He calls "the Newtonian mechanical image . . . unrealistic . . ." (1960, p. 1475).

Halter and Jack take up the subject later (1961), and Back once again offers a discouraging judgment that "none of the particular philosophies of science currently in existence is sufficient as a philosophical guide to research in all the problems of agricultural economics" (1961, p. 908). Kelso reverts when he oversimplifies, saying, "The essence of science is predictability" (1967, p. 224).

Meanwhile, John Brewster, who came to be known as the leading philosopher among economists, wrote and taught about the ancient and modern meaning of science. In lectures at the Graduate School of the U.S. Department of Agriculture, published in part in *A Philosopher Among Economists*, he began with a put-down. In a line paraphrased from the author's memory he said, "I don't want to descend into the snake-pit of what constitutes science." In reality, in his writings he does not dissent from general statements such as that of Allen cited above. His objection is more nearly a corroboration of Back's denial of the usefulness of philosophies of science. Targeting the "body of logic—principles of valid reasoning" that was so popular about mid-century, Brewster insists that much of what we need to know "simply cannot be caught up in any system of logic" (taken from author's notes). He doubts that scientific acquisition of knowledge can be highly systematic.

The subject is reintroduced in literature in presentations Kenneth Farrell, then AAEA president, scheduled in his 1977 program. Moles addresses "The Creation of 'Truth' . . .," and Hartman, "Economics of Science and as Culture." Both stress the cultural element in economic research. Moles: "Science and in specific social science" is "one of the many subcultures in western culture." Further: "Culture may be defined as human knowledge" (p. 919). Knowledge, in turn, is reality, seen as arising in hu-

man experience and not in the rote recording of phenomena (p. 920). Hartman: "The historical faith of science is in the ultimacy of a certain kind of natural order and the effectiveness of the scientific method to discover that order. . . . [But] all qualitative dimensions of human experience were [historically] precluded from scientific understanding. . . . Therein lies the problems of an inadequate cultural tradition" (p. 926). In other words, we can do better.

Strangely, few of the scholars who contemplate the meaning of science take particular note of the necessary process of abstraction. Not many point out that the scientific method inevitably involves abstracting from the complicated universe, reducing it to a scale human beings can deal with. Moles touches on this, as he writes that "any social science approach is used like all other knowledge systems to reduce the complexity of the environment" (p. 921). Breimyer, contributing appropriately to a volume dedicated to Geoffrey Shepherd, more specifically remarks on the abstractions that researchers employ—he calls them "the images of the universe to which agricultural economists address themselves" (1982, p. 57).

It would be hyperbole to credit agricultural economists with having conceptualized the esoteric notion of science and the scientific principle. But it would be inaccurate to the point of disloyalty to fail to acknowledge that agricultural economists have come a long way in achieving an astuteness of understanding of the scientific process to which they give their allegiance.

The Place of Theory in Research

Swanson in 1960 offers salient comments on the scientific concept, but even more pertinent is his interpretation of the relationship of theory to investigation. (Investigation is the original term, eventually replaced by research.) One assumes a theory, he writes, and goes from it to "observations instead of vice versa . . ." (p. 1484). This counsel may now seem axiomatic, but early agricultural economists not only saw no need for theory but were skeptical of theoretical formulations. Bushrod Allin quotes an unidentified outlook economist attending USDA's annual outlook conference, who declared, "We do not engage in theoretical research, ours is the job of analyzing the facts . . ." (p. 409). Even earlier (1940) Theodore Schultz writes that "there are

those among us" (otherwise unidentified) who are alarmed by any mention of "theoretical" and are "filled with sophisticated doubts about that lowly species known as the fact-finding empiricists . . ." (p. 60).

Historically, in 1921 Theodore Macklin recognized a need for an "understanding of the economic forces which underlie the production, marketing and consumption of farm products" (p. 41). E. W. Allen in 1926, however, apparently was the first journal writer to defend categorically the primacy of theory over fact-gathering. His strong language may be appropriate to university courses in research methods today.

My thesis is that a constructive purpose is essential to research; that without it there can be no real research. It is the central idea underlying the method of science, determining all else—the means of approach, the data to be secured, the course of procedure . . . (p. 16).

It is true that science progresses by the determination of facts. . . . But isolated facts, however numerous, do not constitute science. To cite the familiar quotation from Poincaré: "Science is built up with facts as a house is with stones, but a collection of facts is no more science than a heap of stones is a house." And similarly, the indiscriminate accumulation of facts without regard to any central idea or relationship is not in the category of research (p. 17).

Subsequent literature is sprinkled with refinements, such as distinctions between hypothesis and theory. Salter in 1942, after joining others in deploring the practice of compiling "stacks of tables, graphs and maps only to ponder over them unfruitfully," declares forcefully that "hypotheses are a necessary condition to the sorting of facts . . ." (p. 238).

Halter in 1958 chooses to emphasize "presupposition," which "refers to the context in which a true or false proposition is stated." Also, "an absolute presupposition is never subject to the question of truth or falsity as is a proposition" (p. 1871).

In what must be regarded as one of the best journal articles on the role of theory in research, Christensen in 1966 notes that "there has been some misunderstanding . . . regarding the relationship between theory, hypothesis, and model." Some writers, he explains, "would imply that model and theory are one or that hypothesis and theory are the same." He continues, "The hypothesis is a conceptual relationship deduced from the formal theory already established and is a deduction derived from a set of known propositions . . ." (p. 138).

A model, incidentally, is seen by Christensen as "a conceptualization of an abstract system of relationships in terms of something more familiar . . ." (p. 138).

The second stage in the application of theory is to draw on the findings of a research study to revise or refine the preexisting theory. Halter would probably call this the forming of a new proposition. Swanson in 1960 touches on this second stage as well. "Any theory is tentative," he writes, "and subject to further modification with the advent of incompatible evidence. . . ." Moreover, he notes the "integration of probabilistic notions . . ." (p. 1484).

Swanson, however, does not have many companions. It seems surprising that a profession that ploddingly learned to appreciate the place of abstract hypothesis, presupposition, or theory in design of research, has remained less secure in interpreting the findings other than as verification or rejection. There are exceptions. Christensen touches on the second stage, as does Moles. And Halter, as might be expected, offers cogent observations. The scientist, he writes, "realizes that his laws and theories are conjectures and tentative hypotheses and . . . may be rejected as false on the basis of new evidence" (1962, p. 224). Any reformulation comes next. But Halter is one of the few to write in these terms.

Halter almost certainly had picked up Karl Popper's thesis about the falsifiability test, which shows up again in 1989 as Just and Rausser cite Popper's belief in "falsification as the rigorous standard for scientific procedure." They also quote Kuhn, who "found no support" for Popper's thesis. Interestingly, they note without endorsing it that "among economists, McCloskey has advanced the view that economic research is basically essays in persuasion" (p. 1179).

All in all, after three quarters of a century of intrepid research into the economics of agriculture, agricultural economists still show considerable insecurity in the final stage of rejecting or reconstructing received economic theory, or composing a wholly new one.

Taxonomies of Research Technique

Even though they were tentative about articulating the scientific method, the earliest researchers in agricultural economics were quick to develop patterns in research technique. Owing to the profession's obsession with numbers,

during at least the first ten years they gave most of their attention to the collection, analysis, and interpretation of statistical data. Among early concerns were those on how to conduct farm management surveys. Authors wrote about how to make telephonic surveys, for example.

Journal writers have had difficulties of terminology. Those difficulties persist, even complicating the writing of this review article. Economists use the language of research technique, systems of logic and of analysis, philosophy, research tools. Heady refers to the theorist as "tool maker." But he says the agricultural economist is a "tooluser" (1949, p. 837), implying that the agricultural economist does not ordinarily make his own theoretical tools. This viewpoint of Heady is consistent with the first years' engrossment with assembly of data, noted above.

As an arbitrary choice, the more systematic treatments of research technique found in literature will be reported in those terms. There is also a vast literature addressed to method, often mislabeled methodology. It will be noted briefly later.

The philosophy of science and systems of logic show up among the more classically trained agricultural economists, in terms such as induction versus deduction, and positive versus normative science. Conklin in 1947 reminds that "economics originated as a deductive 'science' . . . [consistently with] the Greek belief that all knowledge is derivative of 'pure reason.'" He adds that "inductive methods have come more slowly to economics than to most other branches of academic endeavor" (p. 926). Induction and deduction are rival concepts for Parsons, who sets forth other pairings, such as pure versus applied economics.

Halter, who rarely failed to get into a philosophical fray, refers to "the myth of induction" (1962, p. 223); and he and Jack introduce other terms from Latin and Greek: *a priori* versus *a posteriori*, and analytical versus synthetic. They believe "the key propositions of economic theory should be *a posteriori* synthetic" (p. 95).

Brewster, the acknowledged philosopher of agricultural research, did not fail to remark on deduction and induction in the scientific method. However, he writes that "the hard part of research is . . . discovering a conflict between prevailing theory and exceptional experiences. . . . Deductive development of well-stated hypotheses and testing their probable truth comes in very late stages of our search" (1970, p. 234).

A pairing of concepts that appears occasion-

ally is positive versus normative. In a remark that still surprises, Conklin alleges that "economics began to lose its normative character . . . at the hands of the Mercantilists and the Physiocrats . . ." (1947, p. 926). Inferentially, many early agricultural economists hoped their endeavors could classify as positive, leaving them less subject to challenge in terms of implicit "values."

By any test Glenn Johnson must rank as the foremost thinker in such terms. He elaborates normativism and positivism into pragmatism and variants such as conditional normativism. He writes of "knowledge of values versus prescriptive knowledge" and so on (variously in *Research Methodology for Economists*). In a book reporting a 1985 symposium Tweeten invokes the Johnson credo, dallying with distinctions between normative and positive and then striking comparisons between prescriptive and descriptive (p. 142).

The most structured taxonomies for research technique begin with the monumental study commissioned by the Social Science Research Council referred to above. The report, published in the early year of 1928, touches on virtually all the ideas and terms mentioned thus far in this review: the scientific method (or methods), quantitative versus qualitative, induction versus deduction. The authors classify research technique into five categories: statistical, analogy, case, informal statistical, and experimental.

Leonard Salter, Jr., was a promising young agricultural economist who lost his life in the tragic LaSalle hotel fire in Chicago in June 1946. He still ranks as one of the most systematic thinkers in agricultural economics. In November 1942 he wrote about "cross-sectional and case-grouping procedures" in research, and the taxonomy he set forth in his book, published posthumously in 1948 and reprinted in 1967, features "an outline of inquiry" constituted of problematic situation, formulation of problem, hypothesis, processing of evidence, and a terminal test. The terminal test is as to whether "purposive action is instituted and consonant with consequences" (pp. 68–69). As to research technique (which Salter calls methods), the following are listed: experimental, historical, case, qualitative description, analogical, logical, and statistical (pp. 70–77).

For reasons that are not entirely clear, Salter's field of land economics has lent itself to discourses on research technique in its manifold interpretations. The papers for a symposium, published in a 1966 book edited by W. L. Gib-

son, Jr., R. J. Hildreth, and Gene Wunderlich, essentially captures the thinking of the time.

Several writers have since exhibited their own favorite taxonomies. Alan Randall, addressing resource economics (a successor to land economics in attention getting), reverts to pairings. He recognizes "four major schools of thought" and labels them "institutionalist/land economics; neoclassical/rational planning; public choice/utilitarian; and public choice/individualistic" (1985, p. 1022). Shabman follows by "recasting Randall's schools . . . into two groups: the mainstream, which includes his N/RP, PC/U, and PC/I, and the institutional, or his I/LE." Also, he (Randall) does not "distinguish resource economics from the larger discipline" (p. 1030).

If research technique is the most discretely identifiable structuring of research, ideas surrounding research philosophy are the least so. The term, philosophy, shows up randomly, with the single exception that John Brewster virtually built his career on philosophical concepts. The collection of Brewster's articles, edited by J. Patrick Madden and David E. Brewster, has been noted above. Kenneth Parsons, in 1949, mulls over "what we may call the philosophical frame of reference" (p. 658). Karl Brandt, in 1955, recommends a study of philosophy; and Halter, citing Brandt's faith that it "would lend clarity to our understanding of the nature of economics and to its research methods," adds that "this reference to philosophy has gone unnoticed" (1958, p. 1871). Brewster's star was rising fast at the time, but Halter presumably had not yet taken note of it. As is mentioned above, Halter and Jack later wrote about "a philosophy of science for agricultural economics research." Back entered the fray in 1961, as did a few other economists later. Generally, though, except for the work of Brewster and just a few others, philosophy as orientation to agricultural economics has proved to be almost ephemeral.

The Catch-all of Method—or "Methodology"

A profession that has preferred to minimize the personal and maximize the procedural or instrumental fills its literature with discourses on operating tactics or procedural patterns, usually tagged by the Mother Hubbard term method or, sometimes and erroneously, "methodology."

With regard to the term itself, Don Paarlberg is perhaps the most candid in objecting to methodology, as "this employment of the term is out

of keeping with scientific usage. . . . Most agricultural economists use the word methodology when they are concerned neither with philosophy nor with logic but simply with method. The three syllables are added to convey prestige" (p. 1386).

Frank H. Knight, famed economist of the University of Chicago, told agricultural economists that he finds fault even with addressing method. "I must begin by confessing a degree of skepticism about the practical value of writing on 'method' (a term less pretentious than 'methodology') at a highly abstract level" (p. 112).

But the term method is itself ill defined. Perhaps each method can be viewed as a substratum in a research-technique taxonomy. A great many of the methods that are discussed in literature today relate to the three-century-long preoccupation mentioned at the beginning of this paper, namely, that with numbers.

In a major review published in volume two of *A Survey of Agricultural Economics Literature*, itself anticipated in a 1968 *AJAE* article, George Judge sketches historical works on what he calls estimation in economics. He puts methods of estimation in the category of measurement, and sees "a systematic use of economic and statistical models, methods, and data" as giving "empirical content to economic theory and practice . . ." (p. 3).

As it is questionable whether the prolix and diverse writings about method can be tested against scientific principle, all that will be offered here is a sampling of the practices and conceptualizations that can loosely be called method. Veterans will recognize a number that were prominent briefly. Most of them disappeared. Faddism is not absent in agricultural economics.

It all began with methods of collecting data (the counting era). Next we find a panoply of "statistical methods of analysis" ranging from measures of central tendency to the new simple and multiple correlation, followed by simultaneous equations and other more elaborate forms. The particular equilibrium drew a brief focus of attention and methods were applied to it. In 1953 Fred Waugh, asked to review the applicability of developments in methodology, addressed method instead and listed the following, among others: sampling, design of experiments, consumer panels as data source, structural analysis, economic models generally (they were just coming into popularity), linear programming (due to have a long life), theory of games. In the same

issue of the *Journal*, Richard King introduces activity analysis. Input-output models show up periodically. Marc Nerlove made distributed lags a household word in agricultural economics. The Markov process has been popular intermittently, as has simulation. Dynamic programming, control theory, gini ratios, and a host of other terms and coinages dot the writings in agricultural economics. Some are of lasting merit, others a passing fancy.

In a book reporting a 1985 symposium (published in 1988), Glenn Johnson and Stanley Johnson present a comprehensive inventory of the methods currently at agricultural economists' disposal. Glenn Johnson calls his list technological innovations, and Stanley specifically names quantitative techniques. A discussant, David Bessler, says Stanley surveys "virtually all interesting (useful) techniques in use today" (p. 199).

The Judge inventory of methods in estimation is another useful compendium (1977).

The almost countless designs for conceiving and carrying out agricultural economic inquiries, primarily quantitative in nature and commonly classified under the general term method, are a part of the profession's fabric and yet are peripheral to the most signal tenets of the scientific principle. For this reason they are given only this serial mention.

On Behaving as a Scientist

It ought to have been evident even to the first economists who called themselves agricultural that it is not possible, or even desirable, to carry out agricultural economic studies so impersonally as to neutralize the human factor—converting, in Conklin's language, the investigators into mechanics (1960, p. 1475). The role of the scientist is endemic to all scientific inquiry. As recently as 1989, when the prestigious National Academy of Sciences commissioned a thought-piece on "the nature of scientific research" the opening question, "Is there a scientific method?" was answered in terms that match the thesis of this review, that dedication to scientific principle does not lead to a singular system of scientific practice. It is significant that the resulting publication carries the title, *On Being a Scientist*.

Throughout the history of science, some philosophers and scientists have sought to describe a single systematic method that can be used to generate sci-

entific knowledge. For instance, one school of thought, dating back at least to Francis Bacon . . . , points to observations as the fundamental source of scientific knowledge. . . . By gathering facts without prejudice, a scientist will eventually arrive at the correct theory.

Some scientists may believe in such a picture of themselves and their work, but carrying this approach into practice is impossible. . . . Scientists may be able to suspend some prior theoretical or thematic preconceptions . . . but they cannot view the physical world without any perspective. (p. 2)

In spite of the futility of denying the scientist's role, in the early literature of agricultural economics that element in research is rarely mentioned. Halcrow and Brinegar forthrightly put into words a silently professed thesis of the early years: "Retgression in scientific analysis occurs when allegiance is pledged to men rather than to impersonal processes" (p. 122).

As was predictable, questions of the human factor and economists' performance could not be muffled indefinitely. The moral quality and technical competence of the "men" in agricultural economics drew the attention of a growing number of these individuals.

Interesting and perhaps surprising is Theodore Schultz's quotation from John Maynard Keynes as to the qualities of an economist, who "must reach a high standard in several different directions and must combine talents not often found together." He must be catholic and also "purposeful" but also "disinterested in a simultaneous mood, and as aloof and incorruptible as an artist . . ." (p. 60).

As noted above, Glenn Johnson (in various occasional papers and *Research Methodology for Economists*) led a sizable cadre that matched subjective against positive and introduced terms such as normative and conditionally normative. These relate not so much to impersonal technique as to the role of the human being who collects and examines the data and tries to draw conclusions from them. Manifestly, in this context the researcher is no longer seen as innocuously neutral. Members of the profession listened to Johnson respectfully but did not always confine themselves to his terminology.

Values, Objectivity, Morality, Ethics

Viewing scientific performance in terms not of a formal impersonal system but of the conduct of the scientist opens up questions that the early agricultural economists wanted strenuously to

avoid and that continue to be unattractive to some members of the profession. It suggests that science and practices under its banner are partly a matter of the individual scientist's stance—his outlook, his *Weltanschauung*. A number of agricultural economists have been not only aware of this opposite side of science's coin but have pondered it and written about it. Again, the reference points have varied but have embraced concern for values, objectivity, morality, ethics.

The historical record shows several episodes of preoccupation with values. In 1949 Parsons touched on "the value problem in economic behavior" (p. 679). Frank Knight, reviewing Parsons, writes of "the ambiguous nature and true role of values" (p. 120). He sees two meanings to the word, value. "When used in economic theory, this term refers only to individual or subjective desire, as expressed in choice among alternatives presented" (p. 120). "In contrast," though, "economic policy, and all discussion of any common policy, inevitably refers to . . . some kind of 'objectivity,' some degree of right-and-wrong, capable of being argued." Further, "It involves a distinction between desire and choice as facts versus what they ought to be . . ." (p. 120).

In 1961 Earl Heady as director of the Center for Agricultural and Economic Adjustment at Iowa State University doubtless was instrumental in staging a seminar at that university on "Goals and Values in Agricultural Policy." The papers comprise a book bearing the title.

Many agricultural economists remained diffident about the idea of values, apparently satisfied with vague notions of cultural influences on a researcher. But in 1954 L. John Kutish demonstrated that "the value question" could generate the heat of spirited exchange.

Brewster, as would be expected, touched on values frequently: in "The Cultural Crisis of our Time" (*A Philosopher Among Economists*, pp. 7–65); in "Society Values and Goals in Respect to Agriculture" (*Goals and Values in Agricultural Policy*, pp. 114–37); and in "Beliefs, Values, and Economic Development" (*Journal of Farm Economics*, November 1961).

In all his writings Brewster rarely bothered to define the term, value. He regarded values, redundantly, as "judgments as to what is valuable," and he had society's judgments in mind (p. 158). He implicitly expected the agricultural economist as scientist to accept and be faithful to society's values, and not substitute his own.

Values remain a part of the vocabulary of agricultural economics, but the word has given way

before a related concept that may entail even more sensitive responses, that of objectivity. As Heady suggests, objectivity itself can be regarded as a value; in a seeming contradiction, it adjures the value-steeped economist not only to set aside his values (other than possibly the most elevated ones) as he addresses economic issues, but also to resist all enticements to bias or adulteration.

It may be supposed that agricultural economists generally subscribe, if only subliminally, to the principle of objectivity, yet are not eager for questions to be raised. In 1950 the ever-stimulating Frank Knight, not one of the clan, called attention to the principle in relation to "fact" and "generalization" (p. 113). "Nature," he writes with more than a little whimsy, "is 'honest.'" But "in human conduct" it is necessary to allow for unpredictability as "affected by *error* [his italics]," in several meanings, and by ignorance and prejudice; and where social relations are involved, "outright deception is a fundamental fact" (p. 114).

Harsh language.

When Breimyer in 1967 opened up the subject, he was more guarded or even gentle as he presented his stand that "if agricultural economics is to be a science . . . the precepts of scientific objectivity must be adhered to" (p. 340). He then diplomatically suggests that "the most persistent of all blocks to the objectivity test . . . [may be] the human tendency to avoid areas of controversy . . ." (p. 347). Even so, he does not fail to note "the most recent of all threats . . . [which] arises when an economist divides his time and allegiance so as to multiply his income" (p. 348). And he leaves no doubt that in his mind the test of objectivity must be met if agricultural economics is to "wear the esteemed toga of a science" (p. 350).

Whether the Breimyer literary fist was gloved or mailed, it incited responses. Grove rebuts that "most agricultural economists pretend to an objectivity they do not possess" (p. 153), and the only recourse is subjective dissent (p. 155). Schmitt and Timmerman concur in the judgment of nonattainability of objectivity (p. 921).

Castle, in another follow-up, admits that objectivity is hard to come by and lists a number of threats to it including desire for approval (pp. 810–11).

Alan Randall, the next economist to consider the behavioral paradigm of the scientist, examines "the moral responsibilities of the scientist and his host institution. . . ." He "explicitly acknowledges the role of power in economic interaction." The scientist and his in-

stitution are confronted with "the necessity of making distinct moral choices . . ." (1974, p. 227).

If Randall in 1974 selected the language of morality, in 1983 Maurice Kelso chose yet another variation on the theme as he set up at the AAEA annual meeting a session on "Ethical Issues in Resource Economics" (cf., e.g., the discussion papers by Glenn Johnson and Raymond Anderson).

And so, during forty years of journal exchanges among some of the leading lights in agricultural economics, the bearing of economists' behavior patterns on the scientific quality of the profession's output is examined in variable language but with consistent purpose. And in 1989 Just and Rausser reintroduce Heady's original term of objectivity. They call "insistence on objectivity . . . one of the dominant characteristics of the profession." As did Grove twenty-one years earlier, they see the best protection "in the clash of individual subjectivities" (p. 1179). Only if some individuals are objective and declare themselves can the majority be expected to be so.

The Unsystematic or Uncodified Component to Investigations in Agricultural Economics

In a striking contrast, agricultural economists' testimonials to science and formal scientific practice are mixed with occasional enjoinders to the opposite. Surely, systematic rigor is to be extolled and even mandated. But there is a place also for imagination, inspiration, even perhaps for day-dreaming; and for anecdotal evidence. A number of writers have so declared, sometimes insistently.

Swanson senses the place of intuition. "It is tempting to speculate on the future of the relative emphasis within our profession on the role of the informal predicting expert who uses judgment, intuition and wisdom and that of the researcher methodically seeking to explain phenomena" (p. 1485).

George Ladd carries the argument further, in trenchant language. "We pay little attention . . . to the research tools that are the most versatile and frequently used of all . . . [namely] subconscious mental processes (imagination, intuition, hunch), chance (including serendipity), and writing" (1979, p. 1). In his book, *Imagination in Research*, he explains further that "we need imagination to provide the insights leading to theories that fit the facts. And we rely on in-

tuition for the postulates that support our rules of logic" (p. 134). Ladd even goes so far as to note a place for doubt and tension as contributing to imagination for research.

Brewster, nevertheless, ranks as the foremost exponent of the unsystematic ingredient to research that may otherwise class as scientific:

"Yet . . . there is much to the research quest . . . that cannot be reduced to systematic rules of right reasoning. No rule can be given that will induce one to perceive conflicts between prevailing generalizations and exceptional observations . . . [and] will lead one to induce . . . new hypotheses. . . . In this important sense, research includes far more than mere logic, whether deductive or inductive. It includes insight, genius, groping, pondering—"sense" which can't be boxed up in any formalized procedures. The logic we can teach; the art we cannot . . . (1970, pp. 234–35).

An Epilogue: The Game of Lambasting and Defending Mathematics and Model-building

For longer than a generation, agricultural economists have sparred over what its critics regard as an obsession with mathematics and the constructing of models, often intricate ones. Most of the exchanges have been good-spirited. There is substance to the debates, and a sampling is presented here.

This feature of the historical record is nevertheless presented as an addendum or epilogue because it is not fundamental to an account of the place of science and the scientific principle in agricultural economics.

Most of the verbal tussling has related to the third of the three stages of agricultural economics' preoccupation with numbers, that is, econometrics. During the years of counting, the lightest touch was to admonish field surveyors to phrase their questions carefully. "If the farmer can't answer, look to how you ask the question!"

A few ripostes followed the introduction of the exciting new correlation techniques. Don Paarlberg, for example, says we went crazy and "correlated everything with everything else . . ." (p. 1390).

The early beginning of arguments about the place of mathematics in agricultural economics is attested to by Thompson's pooh-poohing them (1937). "Those who make too strong a claim for the mathematical method appear to believe that this method can of itself yield valid conclusions

. . .” Not so: “there is no *logical* difference between reasoning without mathematics and that which uses higher mathematics. . . .” Then a pox on the other house: “The opposite view that the mathematical method is completely useless for economics usually rests on the mistaken identification of this method with measurement” (p. 718).

Arguments waxed warmer as research moved more fully into econometrics with all its intricate wizardry. In 1961 Karl Fox set up, at the winter association meeting, a session on “Contribution of Econometrics to Farm Price and Income Policy.” Breimyer in a discussion paper offers a judgment consistent with Thompson many years before, one that also anticipates the barrage of evaluations that fill the journals of recent years. “In the magic of electronics the limiting factor is not capacity of computation but expertness of programming. Moreover, neither the electronic machines nor mathematical concepts can compensate for any deficiency in underlying economic theory and understanding . . .” (p. 383).

Paarlberg in 1963 calls it a delusion to think “that the new methods, being mathematical, shelter us from the hazards of human judgment . . .” (p. 1390). Two years later Norris Pritchard pleads for recognizing the limitations of mathematical methods (pp. 152–53). Thereupon Hildreth was moved to ask, at the 1965 annual meeting, “Have we gone too far?” He cites (pp. 1497–98) Paarlberg’s further warning that “no method can produce rational results from erroneous data” (p. 1390) and Kelso’s harsher charge that “the elegance, the tidiness” surrounding mathematical models may only lead us to be “wrong in a more elegant manner” (1965, p. 11). For balance he quotes Waugh’s judgment that “sophisticated techniques are necessary to analyze difficult problems in our complex world” (1964, p. 866). And Heady: quantitative tools have greatly improved the agricultural economist’s ability to handle the masses of data necessary for analysis (1963, pp. 120–22).

In the February 1972 issue of the *Journal*, Oscar Burt and William Martin exchange barbs that are germane to the adequacy-of-data versus ingeniousness-of-technique dispute. Burt had written an article on the economics of investment in specified range improvements. In *Journal* pages the two economists combat as Burt defends his adapting a dynamic programming model to data of acknowledged shortcomings, leading Martin to charge that the exercise amounts

to redefining a problem to fit “the specifications of the tool.” “Elegance,” he added (Kelso’s word again), “is the medium . . .” and has become the message (p. 134). Burt then alleges that Martin confuses “elegance with rigor” (p. 135).

Has much changed since George Warren sought to get better data to put into the farm management studies then being developed? Perhaps not. Warren, however, did not pretend to elegance.

The issue has stayed joined since Warren’s day, yet an emerging consensus may be sensed. If so, it is likely to be consistent with Hildreth’s judgment in 1965 that we have not gone too far in quantitative dexterity (p. 1501) and the real need is “for better training in economics and methodology . . .” (p. 1503).

A consensus judgment may be that employment of mathematical methods does not substitute for rigorous theoretical formulations in agricultural economics. Rather, it itself requires more precise understanding of the constructs of which the discipline is constituted.

Inasmuch as rhetorical exchanges among agricultural economists tend to be cyclical, a new outpouring of challenge and response on the practices attending the econometric stage in agricultural economics can be anticipated in years to come.

All of which introduces a jingle Frederick Waugh attributes to Herman Southworth (1953, p. 706):

Our economic methodology
is full of fine epistemology.
But when we come to problems practical
our theories are too didactical.
If economics is a science,
it needs to foster the alliance
of theorist and statistician,
with manager and prognostician;
To tie the work of mathematicist
to problems of the market strategist.

[Received August 1990; final revision received November 1990.]

References

- Allen, E. W. “Need for Specific Objectives in Economic Research.” *J. Farm Econ.* 8(1926):16–25.
- Allin, Bushrod W. “Theory: Definitions and Purpose.” *J. Farm Econ.* 31(1949):409–17.
- Anderson, Raymond L. “Ethical Issues in Resource Economics: Discussion.” *Amer. J. Agr. Econ.* 65 (1983): 1035–36.
- Back, W. B. “Discussion: The Identification of Problems in Agricultural Economics Research.” *J. Farm Econ.* 42(1963):1471–74.

- . "Philosophy, Method and Status of Agricultural Economics." *J. Farm Econ.* 43(1961):898–909.
- Bessler, David A. "Quantitative Techniques: A Discussion." *Agriculture and Rural Areas Approaching the Twenty-first Century*, ed. R. J. Hildreth et al. Ames: Iowa State University Press, 1988.
- Brandt, Karl. "The Orientation of Agricultural Economists." *J. Farm Econ.* 37(1955):783–806.
- Breimyer, Harold F. "Discussion: Contribution of Econometrics to Farm Price and Income Policy." *J. Farm Econ.* 43(1961):382–85.
- . "The Images Agricultural Economists Think By." *Economic Analysis and Agricultural Policy*, ed. Richard H. Day. Ames: Iowa State University Press, 1982.
- . "The Stern Test of Objectivity for the Useful Science of Agricultural Economics." *J. Farm Econ.* 49(1967):339–50.
- Brewster, John M. *A Philosopher Among Economists*, ed. J. Patrick Madden and David E. Brewster. Philadelphia: J. T. Murphy, 1970.
- . "Beliefs, Values, and Economic Development." *J. Farm Econ.* 43(1961):779–96.
- Burt, Oscar E. "A Dynamic Economic Model of Pasture and Range Investment." *Amer. J. Agr. Econ.* 53(1971):197–205.
- . "More Sophisticated Tools for Less Important Problems: The History of Range Improvement Research: Reply." *Amer. J. Agr. Econ.* 54(1972):134–35.
- Castle, Emery N. "On Scientific Objectivity." *Amer. J. Agr. Econ.* 50(1968):809–14.
- Christensen, Robert L. "Logic, Science, and Economics." *J. Farm Econ.* 48(1966):137–39.
- Conklin, Howard E. "A Neglected Point in the Training of Agricultural Economists." *J. Farm Econ.* 29(1947):925–37.
- . "Are Agricultural Economists Becoming Mechanics?" *J. Farm Econ.* 42(1960):1475–86.
- Ely, Richard T. "American Association for Agricultural Legislation." *J. Farm Econ.* 1(1919):109–14.
- Ezekiel, Mordecai. "On the Use of Partial Correlation in the Analysis of Farm Management Data." *J. Farm Econ.* 5(1923):198–213.
- Gibson, W. L., Jr., R. J. Hildreth, and Gene Wunderlich, eds. *Methods for Land Economics Research*. Lincoln: University of Nebraska Press, 1966.
- Grove, Ernest W. "The Stern Test of Objectivity for the Useful Science of Agricultural Economics: Comment." *J. Farm Econ.* 49(1967):153–55.
- Halcrow, Harold G., and George K. Brinegar. "A Brief Note on Schools of Thought." *J. Farm Econ.* 35(1953):122.
- Halter, A. N. "A Metaphysical Hypothesis." *J. Farm Econ.* 40(1958):1871–74.
- . "Reply to 'Philosophy, Method and Status of Agricultural Economics.'" *J. Farm Econ.* 44(1962):221–25.
- . "The Identification of Problems in Agricultural Economics Research." *J. Farm Econ.* 42(1960):1459–71.
- Halter, A. N., and J. J. Jack. "Toward a Philosophy of Science for Agricultural Economic Research." *J. Farm Econ.* 43(1961):83–95.
- Hartman, L. M. "Economics as Science and as Culture." *Amer. J. Agr. Econ.* 59(1977):925–30.
- Heady, Earl O. "Implications of Particular Economics in Agricultural Economics Methodology." *J. Farm Econ.* 31(1949):837–50.
- . "The Agricultural Economist and His Tools: 1. Research Methods." *Proceedings of Eleventh International Conference of Agricultural Economists*. London: Oxford University Press, 1963.
- Hildreth, R. James. "Have We Gone Too Far?" *J. Farm Econ.* 47(1965):1497–1503.
- Iowa State University Center for Agricultural and Economic Adjustment. *Goals and Values in Agricultural Policy*. Ames: Iowa State University Press, 1961.
- Johnson, Glenn L. "Ethical Issues in Resource Economics: Discussion." *Amer. J. Agr. Econ.* 65(1983):1033–34.
- . *Research Methodology for Economists: Philosophy and Practice*. New York: Macmillan Co., 1986.
- . "Technological Innovations with Implications for Agricultural Economics." *Agriculture and Rural Areas Approaching the Twenty-first Century*, ed. R. J. Hildreth et al. Ames: Iowa State University Press, 1988.
- Johnson, S. R. "Quantitative Techniques." *Agriculture and Rural Areas Approaching the Twenty-first Century*, ed. R. J. Hildreth et al. Ames: Iowa State University Press, 1988.
- Judge, George G. "The Search for Quantitative Economic Knowledge." *Amer. J. Agr. Econ.* 50(1968):1703–17.
- . "Estimation and Statistical Inference in Economics." *A Survey of Agricultural Economics Literature*, vol. 2, ed. Lee R. Martin. Minneapolis: University of Minnesota Press, 1977.
- Just, Richard E., and Gordon C. Rausser. "An Assessment of the Agricultural Economics Profession." *Amer. J. Agr. Econ.* 71(1989):1177–90.
- Kelso, M. M. "A Critical Appraisal of Agricultural Economics in the Mid-Sixties." *J. Farm Econ.* 47(1965):1–16.
- . "Definitions of Science: Reply." *J. Farm Econ.* 49(1967):224–25.
- King, Richard A. "Some Applications of Activity Analysis in Agricultural Economics." *J. Farm. Econ.* 35(1953):823–33.
- Knight, Frank H. "Comment on Professor Parsons' Article." *J. Farm Econ.* 32(1950):112–22.
- Kuhn, T. *The Structure of Scientific Revolutions*, 2nd ed. Chicago: University of Chicago Press, 1970.
- Kutish, L. John. "The Value Question." *J. Farm Econ.* 36(1954):666–71.
- Ladd, George W. "Artistic Tools for Scientific Minds." *Amer. J. Agr. Econ.* 61(1979):1–11.
- . *Imagination in Research: An Economist's View*. Ames: Iowa State University Press, 1987.
- Macklin, Theodore. "Report of Committee on Farm Economic Investigational Work, the American Farm Economic Association." *J. Farm Econ.* 3(1921):41–44.
- Martin, William E. "More Sophisticated Tools for Less Important Problems: The History of Range Improvement Research: A Comment." *Amer. J. Agr. Econ.* 54(1972):132–34.
- McCloskey, F. *The Rhetoric of Economics*. Madison: University of Wisconsin Press, 1985.

- Moles, Jerry A. "The Creation of 'Truth' by Social Scientists and Planners: Assumptions, Decisions, and the Negotiation of Reality." *Amer. J. Agr. Econ.* 59(1977):918-24.
- National Academy of Sciences, Committee on the Conduct of Science. *On Being a Scientist*. Washington DC: National Academy Press, 1989.
- Nerlove, Marc. "Distributed Lags and Estimation of Long-run Supply and Demand Elasticities: Theoretical Considerations." *J. Farm Econ.* 40(1958):301-13.
- Paarlberg, Don. "Methodology for What?" *J. Farm Econ.* 45(1963):1386-92.
- Parsons, Kenneth H. "The Logical Foundations of Economic Research." *J. Farm Econ.* 31(1949):656-86.
- Popper, Karl R. *The Logic of Scientific Discovery*. New York: Basic Books, 1959.
- Pritchard, Norris T. "Limitations of Mathematical Methods." *J. Farm Econ.* 47(1965):152-53.
- Randall, Alan. "Information, Power and Academic Responsibility." *Amer. J. Agr. Econ.* 56(1974):227-34.
- . "Methodology, Ideology, and the Economics of Policy: Why Resource Economists Disagree." *Amer. J. Agr. Econ.* 68(1985):1022-29.
- Salter, Leonard A., Jr. *A Critical Review of Research in Land Economics*. Madison: University of Wisconsin Press, 1967.
- . "Cross-sectional and Case-Grouping Procedures in Research Analysis." *J. Farm Econ.* 24(1942a):792-805.
- . "The Content of Land Economics and Research Methods Adapted to its Needs." *J. Farm Econ.* 24(1942b):226-47.
- Schmitt, G., and W. Timmermann. "On Scientific Objectivity." *Amer. J. Agr. Econ.* 51(1969):921-24.
- Schultz, Henry. "The Statistical Measurement of Elasticity of Demand for Beef." *J. Farm Econ.* 6(1924):254-78.
- Schultz, Theodore W. "Needed Additions to the Theoretical Equipment of an Agricultural Economist." *J. Farm Econ.* 22(1940):60-66.
- Shabman, Leonard. "Natural Resource Economics: Methodological Orientations and Policy Effectiveness." *Amer. J. Agr. Econ.* 67(1985):1030-38.
- Shepherd, Geoffrey S. *Agricultural Price Analysis*, 2d ed. Ames: Iowa State College Press, 1941.
- . "What Can a Research Man Say About Values?" *J. Farm Econ.* 38(1956):8-16.
- Social Science Research Council. *Research Method and Procedure in Agricultural Economics*. Mimeographed. Washington DC, 1928.
- Spillman, W. J. "Remarks by W. J. Spillman." *J. Farm Econ.* 4(1922):99-100.
- Swanson, Earl R. "Discussion: Are Agricultural Economists Becoming Mechanics?" *J. Farm Econ.* 42(1950):1483-86.
- Taylor, Henry C., and Anne Dewees Taylor. *The Story of Agricultural Economics in the United States, 1840-1932*. Ames: Iowa State College Press, 1952.
- Thompson, J. M. "Mathematics in Economics." *J. Farm Econ.* 19(1937):718-26.
- Thurman, Walter N., and Mark E. Fisher. "Chickens, Eggs, and Causality, or Which Came First?" *Amer. J. Agr. Econ.* 70(1988):237-38.
- Tweeten, Luther. "Domestic Food and Agricultural Policy Research Directions." *Agricultural and Rural Areas Approaching the Twenty-first Century*, ed. R. J. Hildreth et al. Ames: Iowa State University Press, 1988.
- Waugh, Frederick V. "Applicability of Recent Developments in Methodology to Agricultural Economics." *J. Farm Econ.* 35(1953):692-706.
- . "What Do Our Readers Want?" *J. Farm Econ.* 46(1964):866.
- Working, Elmer J. "What Do Statistical Demand Curves Show?" *Quart. J. Econ.* 41(1927):212-27.

Rent Seeking: The Potash Dispute between Canada and the United States

Valerie J. Picketts, Andrew Schmitz, and Troy G. Schmitz

Despite the Canada-United States Free Trade Agreement, border disputes involving agricultural trade are common. A theoretical basis for the 1987 U.S. countervailing duty case against Canadian exports of potash is developed using excess capacity arguments. Empirically, within a rent-seeking context, the U.S. potash producers gained far less from their legal action against Canada than did U.S. farmers who are significant users of Canadian potash.

Key words: agricultural trade, countervailing duties, potash, producer rents, rent seeking, user rents.

Border disputes involving trade in agricultural inputs between Canada and the United States have received relatively little attention in the economic literature even though the problem is of major significance. Charges of export dumping continue despite the signing of the Canada-United States Free Trade Agreement. This paper focuses on countervailing duty actions against Canadian exporters of potash to the United States, which is a major user. Potash production in Canada is a multimillion dollar business as it is both the major supplier of potash to the U.S. market and the largest exporter of potash in the world.

The structure of the potash industry in North America is oligopolistic because of the existence of a few players and the ability of the larger firms to affect prices (Anderson, Bennett, Fulton and Schmitz, Olewiler, Flatters and Olewiler). Two such firms are the International Mineral and Chemical Corporation (IMCC) and the Potash Corporation of Saskatchewan (PCS), which control almost half of the potash production in North America. The largest player in the North American potash market is Canada, where

Saskatchewan accounts for approximately 95% of Canadian production and New Brunswick produces the other 5%.¹ Canadian production accounts for approximately three-quarters of North American production (Korol, p. 6). Potash is sold both to the United States and in offshore markets. Canpotex, Ltd., sells all of the exports of the Saskatchewan potash producers to the offshore market.² In contrast, each of the several Saskatchewan producers market potash independently in the U.S. market. Approximately two-thirds of Saskatchewan production is sold to the U.S. market, and the remaining one-third is sold offshore (Korol, p. 9).

Most of the 1970s and early 1980s were prosperous times for the potash industry in North America, largely because of high grain prices and high farm incomes. However, the potash market changed quickly as a result of falling grain prices; and, by the middle of the 1980s, potash producers in North America had experienced substantial losses. This prompted two American potash producing firms in 1987 to sue the Canadian potash producers for dumping potash on

Valerie J. Picketts is livestock market analyst at the Economics Branch, Saskatchewan Agriculture and Food; Andrew Schmitz is chairman and a professor, Department of Agricultural and Resource Economics, University of California, Berkeley; Troy Schmitz is a research assistant, Department of Agricultural and Resource Economics, University of California, Berkeley.

The authors thank the three anonymous referees and Casey Van Kooten for insightful comments.

¹ Saskatchewan produces muriate of potash (potassium chloride), which is one form of potash. Over 95% of North American production is muriate of potash, and it is considered a homogenous good. Approximately 95% of the muriate of potash produced in North America is used as fertilizer, and there is no substitute for it in the agricultural market (Korol, p. 6). Muriate of potash is generally not used on Canadian soils; therefore, U.S. farmers are the major consumers of Canadian potash in North America.

² New Brunswick opted to stay out of Canpotex, Ltd., which is a company that sells Saskatchewan potash offshore.

the U.S. market (Bowley). Countervailing duties were then placed on Canadian potash exports entering the United States. Subsequently, the Canadian potash firms increased their prices; in response, the duties were lifted on Canadian potash.

The legal action of the U.S. potash firms can be viewed as rent seeking along the lines presented in an earlier paper by Bredahl, Schmitz, and Hillman on the tomato war between the United States and Mexico. An intensive campaign was launched by U.S. potash producers in lobbying the U.S. government into imposing duties on potash imports (Marud). The duties created economic rents for the U.S. potash producers by increasing the price received. However, U.S. farmers lost rents as a result of the price increase. They attempted to prevent the imposition of the duties but failed. In contrast, Canada incurred a net gain in rents resulting from the price increase.

The potash dumping case is interesting because it appears that rent seeking on the part of U.S. potash firms created sizable gains for Canada at the expense of U.S. farmers. Often, in agricultural border disputes, the opposite is the case. Producers of commodities bring antidumping charges against exporters because dumping behavior depresses commodity prices. The action by U.S. potash producers essentially stopped the Canadian potash producers from pricing below the cost of production (Greenaway). The increase in price caused profits to increase for potash producers in both the United States and Canada. The legal action by the U.S. government, in essence, created a cooperative pricing solution where the rents of both parties increased.

The overall objective of this paper is to formulate a model for analyzing countervailing duties in a mineral market such as potash. The empirical objectives of this paper are to estimate (a) the effects on both Canadian and U.S. potash producers and on U.S. farmers from the implementation of the potash duties in 1987; (b) the rents under two different assumptions on pricing behavior by Canadian producers, namely, (i) dominant market strategy and (ii) competitive market strategy. The rents estimated in (a) above are then compared to the rents generated under the dominant and competitive strategies.

In earlier literature, dumping is defined as price discrimination between national markets. It is thus treated as profit maximization by a discriminating monopolist or as an oligopolistic tactic to eliminate competition or to enforce a cartel.

However, in this paper, we use the modern approach to dumping, where adjustment costs and excess capacity are considered (Ethier). Dumping occurred in the potash industry largely because of excess industry capacity created by falling demand. In this connection, dumping was determined on the basis of "fair market price." This method is often used in antidumping disputes (Schmitz, Firsch, and Hillman). Concerning dumping in raw materials such as potash, Lloyd noted earlier that dumping is particularly likely in industries where there is surplus capacity, either because of cyclical reduction in demand or because of irregular investment patterns in large plants.

The Potash Trade Dispute between Canada and the United States

In 1982, low grain prices and high real interest rates caused potash demand to fall. This led to a large decrease in the production of potash in Canada (Fulton and Schmitz, p. 32). However, U.S. potash production increased because of the production cutback by Saskatchewan producers in their attempt to maintain prices. Despite the cutback, the price fell from \$146 in 1981 to \$109 in 1982 (U.S. dollars per metric ton of potassium oxide; unless stated otherwise, all dollars are in U.S. currency). In 1983 and 1984, potash demand recovered to near previous levels, and Saskatchewan producers appeared to switch to a competitive strategy by decreasing prices in order to gain market share. This resulted in a considerable drop in price, and the Saskatchewan producers' market share and capacity utilization rate increased during 1983 and 1984. In 1983 the consumption increased but the price fell from the 1982 level by an additional \$13 per metric ton. The Canadian market share increased from 82% of agricultural sales in 1982 to 89% in 1984. In 1985 and 1986, the price fell further as a result of a competitive strategy of increasing market share at the expense of price (Korol, pp. 34, 35).

On 10 February 1987, two American firms (Lundberg Industries of Dallas, Texas, and New Mexico Potash Corporation of Memphis, Tennessee) filed a lawsuit against several Canadian producers for dumping potash in the United States at prices which were alleged to be 43% below the cost of production. In 1987 these two firms accounted for less than 15% of the potash consumed in the United States. The suit was filed through the U.S. Department of Commerce and

the U.S. International Trade Commission. The American potash companies perceived that the Canadian producers were price cutting to increase their market share in the United States; the Canadian market share of agricultural sales increased from 82% in 1982 to 88% in 1987 (U.S. Department of Interior). The International Trade Commission agreed on 3 April 1987, that there was indication of "unfair price discrimination"; and on 21 August 1987, the United States announced the preliminary duties on Canadian potash.

The Provincial Government of Saskatchewan responded by limiting potash production in order to increase the world price of potash. If the Saskatchewan potash producers exceeded their production limits, they faced a fine of \$1 million by the provincial government. If they continued to produce after the limit was exceeded, then the fine would be \$50,000 per day. In September 1987, PCS announced that it was increasing its price by 60%—from \$58 to \$93 per metric ton—with the expectation that the other Saskatchewan producers would follow its lead. Following the lead of PCS were IMCC; Kalium Chemicals; Cominco, Ltd.; and Noranda, Inc. Therefore, PCS successfully acted as a price leader.

The U.S. farmers who stood to lose from the rent seeking of the U.S. potash producers lobbied the U.S. government through organizations such as the National Corn Growers Association. However, the farmers' lobbying was unsuccessful.

In December 1987, the Saskatchewan producers agreed to stop this dumping behavior. As a result, on 8 January 1988, the Canada-U.S. dispute case was suspended for five years by an agreement between the Saskatchewan producers and the U.S. Department of Commerce. The duties on potash were to be removed as long as Saskatchewan firms did not sell their product to U.S. producers below the "minimum price" calculated in accordance with the agreement within the five-year period. The agreement defined the "minimum price" for each producer as the price not less than a price equal to the current "fair value" of its potash minus 15% of its preliminary dumping margin. The current "fair value" of potash for a producer would be either: (a) the constructed value of its potash if it was selling below the cost of production, (b) its Canadian price if it had 5% or more of total sales in Canada and if its price was above its production costs, or (c) its offshore price if it was selling at prices above its cost of production but had an insuf-

ficient volume of sales in Canada. On 25 January 1988, PCS announced that it would decrease its price by 15% of its dumping margin value according to the fair market value definition.

The Model

The potash market is depicted in figure 1. The marginal cost curves are constant over a wide range of production until the capacity constraint is reached, at which point they increase sharply (Bennett, Korol). Also, production costs in the United States are significantly higher than in Canada (Energy, Mines and Resources). In figure 1, the marginal cost curves for Canada and the United States are S^cbs^c and S^us , respectively. These reflect the cost differences between the two countries including transportation costs. They also show that the capacity level of the U.S. firms is relatively small compared to the Canadian firms. The aggregate supply schedule is S^cbedS^* , the demand for Canadian potash in the offshore markets is D_{os} , the demand from the U.S. market is D_{us} , and the sum of the two demands is represented by TD . There is no significant Canadian demand for potash. Hence, a demand schedule for Canada is not drawn. At price P_s , aggregate output is Q_2 , of which Q_1Q_2 is shipped offshore. U.S. output is Q^u , and Canadian output is Q_1 . At the competitive equilibrium price P_s , firms in the United States earn rents equal to the cross-hatched area $ac'de$, while Canadian firms earn rents totaling P_sabS^c . In this model, rents are positive for firms in both countries because aggregate demand TD lies to the right of the capacity constraint demand TD' . Thus, sizable rents can exist even in the absence of some form of cooperative pricing strategy among firms.

If the U.S. and Canadian firms were able to cartelize, the industry output would be restricted and profits would increase. One such example of cooperation would be dominant firm pricing strategy by Canada. Canadian output would be restricted to Q^* , price would increase to P^* , and U.S. output would be unchanged. In this case the largest percentage of profits from cooperation would go to the United States (profits increase by $gfc'a$) along with an increase in the U.S. market share. There is also an optimal import-export cartel pricing solution, different from the above, where market shares between the United States and Canada could remain virtually unchanged. This solution was derived by Bre-

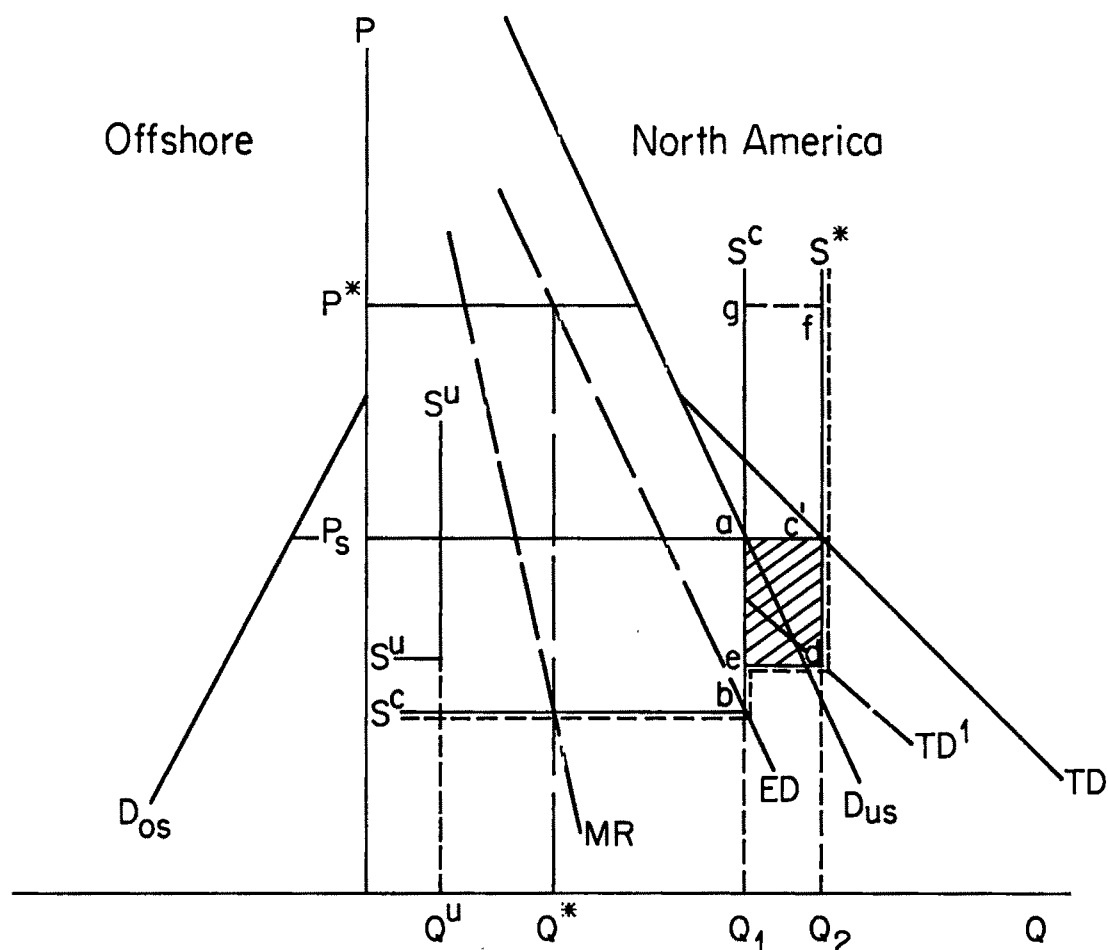


Figure 1. Price equilibrium and dominant firm pricing

dahl, Schmitz, and Hillman in their discussion of the tomato war between the United States and Mexico.

The U.S. dumping litigation against Canadian firms has been assumed away in figure 1 because it does not allow for industry losses. Often the potash industry has excess capacity and loses money (negative rents). The negative rent model in figure 2 depicts the case that existed prior to the introduction of the U.S. dumping lawsuit in 1987. $S^u S^c$ represents the U.S. supply curve, and $S^c S^c$ represents the supply schedule for Canadian firms. Total supply is given by the $S^* d a e S^c$. Given total demand TD , aggregate output is Q_2 at price P_s .

Suppose aggregate demand shifts to TD' . If both countries continue to produce aggregate output Q_2 , both suffer losses. The loss to Canadian firms is $P_1 b e S^c$, while the loss suffered by U.S. firms is $a d f b$. The combined loss is the

cross-hatched area $S^c e a d f P_1$. Note, however, that the U.S. users of potash gain from these over-production activities.

The model in figure 2, where price is P_1 and Canadian output is Q_1 , is clearly a classic dumping case viewed from the standpoint of U.S. potash producers since Canada is selling to the U.S. market below the cost of production. What if the United States is successful in imposing a duty of T , requiring Canada to price at P_0 if they want the U.S. to remove the duty? It is in the best interest of Canadian firms to cooperate and price at P_s in order for the United States to remove its duties with the result that the Canadian firms no longer lose money. However, in equilibrium, if the U.S. firms are not willing to cut output, Canadian firms have excess capacity of the amount $a a'$. Thus, they will have to cut output to Q^* but in so doing will lose market share.

Even with the cutback in Canadian production

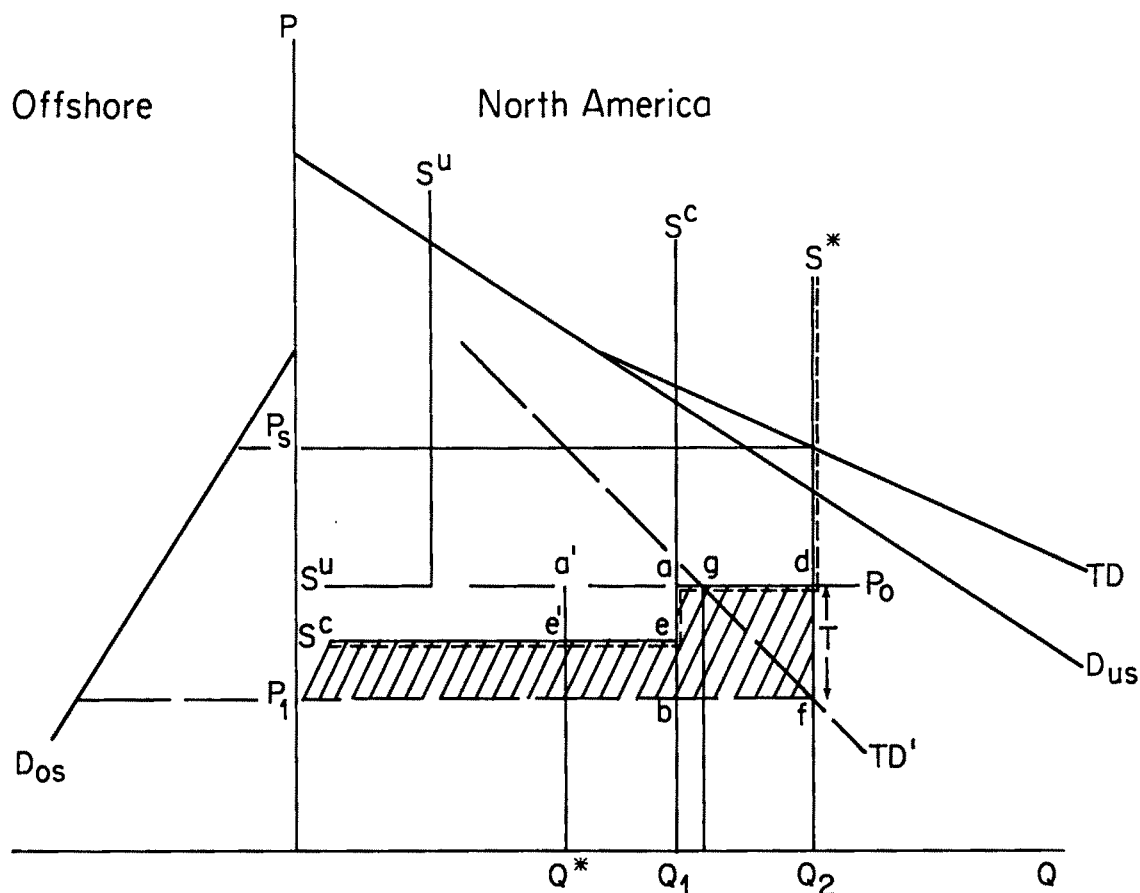


Figure 2. Negative rents in the potash market

to Q^* , there are also gains to Canadian firms. These firms now have positive rents of $S^u a' e' S^c$. However, U.S. potash users lose $S^u g f P_1$. Interestingly, both the U.S. and Canadian industries gain from the temporary duty T if these duties result in a permanent price of P_0 (at which time they can be removed). However, the loss to U.S. potash users is greater than the gain to U.S. potash producers. Clearly, the dumping suit was against the best interest of U.S. corn producers.³

In the model by Bredahl, Schmitz, and Hillman on rent seeking, attempts on the part of U.S. tomato growers at imposing tariffs were viewed as a competitive strategy because, if imposed, the rents to Florida producers would increase while rents to Mexican producers would de-

crease. This, however, assumes that the tariff would be permanent. In the potash case this was not so, as the exporters agreed to raise the price so the tariff could be removed. As a result, the dumping charge by the U.S. producers ended by having the U.S. and Canadian firms cooperating, whereby rents to both were increased as a result of the imposition of duties. However, market share was lost by Canadian firms. A more optimal solution for Canada was clearly one where both the United States and Canada reduced output. Regardless, rent-seeking activities on the part of the U.S. potash producers to persuade Washington to impose temporary duties were, as shown later, costly to U.S. farmers; also, U.S. potash producers gained less than users lost.

Antidumping behavior changes the mean price of fertilizer purchased by U.S. farmers. If demand fluctuates between TD' and TD (figure 2), farmers would pay prices ranging between P and P_s . However, with the threat of antidumping ac-

³ Because of the cost differences between the United States and Canada, if the U.S. insists that Canada does not price below P_0 (using as a threat the reintroduction of countervailing duties), Canadian firms are always guaranteed a positive rent because price exceeds marginal cost.

tivity, farmers pay prices ranging between P_0 and P_s . As a result, the mean price of fertilizer facing U.S. farmers increases.

It appears that a U.S. lawsuit was necessary to change the pricing and production strategy of Canadian producers. Why this apparent irrational behavior by Canadian firms prior to the antidumping litigation? There are several explanations: (a) Canadian potash producers might have expected that potash demand would be weak for only a short period of time (in this case, a strategy of shutting down and reopening several potash mines may have been viewed to be a more costly alternative than that which was actually followed); (b) the producers might not have perceived the magnitude of the demand shifts; (c) there may have been predatory pricing behavior by one or two of the potash corporations within the Canadian complex; and (d) for many potash firms, it is difficult to cut back production drastically in a short period of time. (Ethier). To do this requires substantial layoffs. This is politically unacceptable, especially for PCS which, at the time of the countervailing duty action, was entirely a Crown Corporation. Worker layoffs could be justified much more easily as a response to countervailing duty action.

Regardless of the reasons, the Canadian potash industry was unable to cut back production to raise prices significantly to cover production costs. The U.S. countervailing duty case against potash exports, if not caused, at least coincided with both a reduction in Canadian output and a price increase.⁴ It is perhaps ironic that the Canadian Potash industry was the subject of an antidumping lawsuit. Saskatchewan producers, although they are of a different nature than those that made up the industry in the 1980s, had witnessed a U.S. dumping charge previously. In the 1960s, because of expanding capacity in Canada and the USSR, potash prices fell approximately 50% from 1965 to 1969. In the late 1960s, the economic hardship encountered by the higher-cost New Mexico producers caused layoffs of workers. As a result, political pressure developed to decrease the amount of potash

entering the United States. During the 1960s, Saskatchewan accounted for 93% of U.S. imports, which was 70% of Saskatchewan's production. The U.S. Treasury initiated charges of dumping against Canada. In August of 1969, the U.S. treasury deemed that dumping had occurred and referred the case to the U.S. Tariff Commission to determine the restrictions to be placed on potash entering the U.S. market. At the same time, the Saskatchewan government developed a prorationing scheme to increase the price of potash sold in the United States. Each Saskatchewan producer was granted a quota based on its proportion of total industry capacity. Also, a minimum price was established which, if violated, would lead to the cancellation of the operators' production license. The price set by Saskatchewan firms became the world price, and New Mexico producers were able to operate close to capacity.

The prorationing of the 1970s has made Saskatchewan firms aware of the benefits of co-operating. It caused the price of potash to increase by \$30 per tonne (constant 1980 \$Canadian) from 1969 to 1970. The total quantity produced in Saskatchewan decreased, but their profits increased because of the inelastic nature of potash demand. In 1973 there was a strong increase in demand for potash, and the program was abandoned in 1974 (Anderson, pp. 5-8).

Empirical Analysis

In calculating the rents lost and gained by the various groups from the U.S. antidumping case in 1987, the prices used were free on board (FOB) the potash mine. The prices were deflated to the 1982 level; therefore, all prices and welfare measures are given in 1982 dollars.

Olewiler (pp. 8, 9) estimated a series of average cost curves for various Saskatchewan mines and found that they were U-shaped with a relatively large region of constant unit cost. The cost included all mine and mill operating costs excluding taxation and depreciation. If the mines and mills were operated at full capacity (500,000 to 800,000 metric tons per year), then the average costs would be between \$30 and \$48 per metric ton of potassium oxide, or \$49 to \$79 per metric ton of potassium chloride (in 1982 dollars). In this analysis, the marginal cost used for Canada was based on an average of Olewiler's estimates, and the U.S. marginal cost was assumed to be 57% greater than Canada's mar-

⁴ A shortcoming of our analysis is that one cannot conclude that the price of potash increased as a result of the lawsuit. Perhaps PCS and others would have raised prices anyway. In conjunction with the countervailing duty case, the dynamics of firm behavior in cutting back output takes on many dimensions. PCS appeared to be the leader in raising prices and cutting output; other firms followed. However, for a time at least, PCS lost market share to some of its Canadian competitors. The pricing strategy followed by Canadian potash firms over time is a topic of further research and is not dealt with in this paper.

ginal cost. The estimates of the marginal cost, excluding transportation and retail costs, were \$40 for Canada and \$62 per metric ton of potassium oxide for the United States.

The transportation costs incurred by Saskatchewan and U.S. producers are approximately the same because most U.S. producers are located in New Mexico and, therefore, Saskatchewan and U.S. producers have similar distances to the major domestic market area in the Corn Belt region of the United States (the mid-western United States). The transportation costs will shift back and forth in favor of the Saskatchewan or U.S. producers, depending on the current transportation and economic policies as well as the market conditions. Generally, the northern states obtain most of the potash they consume from Canada, while the southern states are supplied by the Carlsbad, New Mexico, producers. The Saskatchewan and New Mexico producers compete mainly in the central states (Bennett, p. 43). In this paper, the transportation costs for the United States and Canada were assumed to be the same. Transportation costs from Saskatchewan to the major U.S. markets vary between \$42 and \$62 per metric ton of potassium chloride (1983–84 dollars), depending on the type of transportation. An average of \$31 per metric ton of potassium oxide was used for both Canada and the United States because they incur similar transportation costs. Transportation costs account for 25%–50% of the U.S. farmers' average retail price of potash. Fixed transportation and retail costs were included to determine the price paid by farmers.

The demand curve for fertilizer is generally considered price inelastic because there are few substitutes. The demand for potash is derived from the demand for the crops for which potash is an input (Fulton and Schmitz, p. 3). Bennett estimated the price elasticity in the U.S. market to be -0.38 from 1960 to 1980. Heady and Yeh found a price elasticity of demand in the United States to be -0.403 for the time period from 1926 to 1956. Hee estimated an elasticity of -0.46 from 1956 to 1965. Marhatta estimated a price elasticity of demand for aggregate U.S. potash for the period, 1950 to 1974, to be -0.53 . Olewiler (p. 12) estimated a long-run price elasticity of demand in North America from 1973 to 1984 to be -0.36 . Thus, an elasticity of -0.4 was used in the model.

In deriving the demand equations to calculate the user rents, the demand intercept was assumed to be changing while the slope remained constant. Therefore, the assumption was that the

demand was shifting from one time period to the next as a result of possible changes in the prices of corn and other grains that would affect the demand for potash. Allowing the intercept to shift is a more realistic approach than allowing the slope to change which would imply that the farmer response to a change in the price of potash was changing over time. This is illustrated in figure 3. If the demand curve for potash fertilizer shifts from D_1 to D_2 with a corresponding price increase from P_1 to P_2 , then user surplus has increased ($a_2c_2P_2 - a_1b_1P_1$) even though the price has increased. To quantify the impact of the countervailing duty action by the United States, the change in user surplus has to be estimated for a given demand curve. For example, if demand is D_2 and the price is increased from P_1 to P_2 due to legal action, then the change in user surplus is $P_2c_2b_2P_1$. From figure 3, in order to estimate the effects of the countervailing duty case, one has to compare periods before and after the litigation where demands are comparable.

The capacity of the U.S. potash producers was estimated using Korol's producer capacity utilization rates. The maximum value of the U.S. capacity estimated was 834,000 metric tons per six-month period based on a capacity utilization of 93% of production in 1979. The capacity level was assumed to remain constant over the data range from 1982 to 1988.

The rents are calculated for periods before and after the imposition of the tariff; the tariff level is based on the margin percentages imposed on the Saskatchewan firms when the duties were in place.

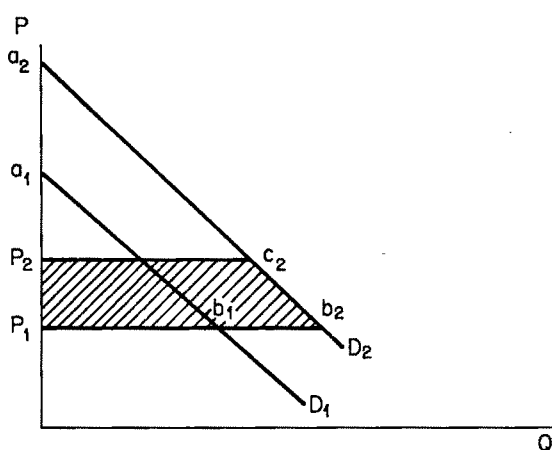


Figure 3. Measure of change in user surplus given a change in demand

The margin percentages for firms in Saskatchewan were:

Firm	Margin (%)
Potash Corporation of Saskatchewan	51.90
International Mineral and Chemical Corporation	9.14
Kalium Chemicals	29.67
Noranda, Inc.	85.20
Potash Company of America	77.40
All others	36.62

A weighted average margin, 41.43%, was used.

The estimated user and producer rents from 1982 to 1988 are given in table 1. In 1982, there was a large decrease in producer rents as prices and sales of potash declined sharply. Producer rents in the United States became negative in 1985 and remained so until 1988. Canadian rents became negative in 1986 and remained so until 1988. The Canadian marginal cost is lower than the U.S. marginal cost and, therefore, Canada did not incur a loss in surplus until 1986 (2).⁵ The large Canadian loss in 1987 (2) from sales in the United States was caused, in part, by the

tariff duties raising the Canadian marginal cost by 15% from 21 August 1987 to 8 January 1988.

The numbers show that after the imposition of the duties, U.S. producer rents were at least \$15 million higher in 1988 (2) than in 1986 (2).⁶ The corresponding number from the period 1986 (1) to 1988 (1) was roughly \$13 million.

As a result of the price increase and the elimination of the tariff duties, Canadian potash producer rents in 1988 (2) were roughly \$175 million higher than in 1987 (1). The gains from increased sales to the United States market exceeded the gains from offshore sales.⁷

The user rents fluctuate from year to year due

⁶ As the theory suggests, for Canada to maintain prices to avoid further countervailing duty action by the United States, Canadian firms have to adjust by cutting back production. As a result, in the North American market, U.S. producers' share would increase. This appears to be what has happened, and our estimates take market share changes into account. For the marketing year ending 30 June 1990, U.S. producers provided roughly one million metric tons to the U.S. market, which is greater than 20% of the market supplied by all North American producers. Prior to the imposition of the countervailing duties, this market share was generally 15% or less.

⁷ Offshore prices also rose and are positively correlated with U.S. prices. Because Canpotex's offshore prices follow U.S. prices with a lag, the offshore rent of \$55.4 million for 1987 is overstated (table 1).

⁵ Throughout the text, (1) refers to the period January–June and (2) refers to the period July–December.

Table 1. Potash: User and Producer Rents, 1982–88

Time Period ^a	U.S. User Surplus	Producer Surplus		U.S.
		Canada		
		To U.S. ^b	To Offshore ^c	
(U.S. \$ thou.)				
Prior to duty ^d				
1982 (1)	364,550	75,027	32,944	11,419
1982 (2)	271,333	47,053	21,204	1,745
1983 (1)	395,411	44,645	25,430	1,415
1983 (2)	559,877	52,726	25,092	1,477
1984 (1)	430,164	61,452	41,837	4,558
1984 (2)	464,418	63,069	33,449	390
1985 (1)	513,917	40,979	18,368	-744
1985 (2)	324,040	14,723	7,869	-3,971
1986 (1)	517,223	12,962	6,453	-5,386
1986 (2)	245,306	-1,074	-850	-8,219
1987 (1)	243,030	-3,842	-2,904	-643
Interim duty ^e				
1987 (2)	298,673	-26,388	25,288	-901
Post duty ^f				
1988 (1)	443,478	71,543	55,431	7,526
1988 (2)	443,898	99,951	70,335	7,234

Source: Calculated.

Note: Rents are based on the actual prices and quantities, 1982–88; elasticity of demand = -0.4

^a Prior to duty, 1982 (1) through 1987 (2), and interim duty, 1987 (2); figures represent United States and offshore.

^b Canadian producer rents from sales to the U.S. market.

^c Canadian producer rents from sales to the offshore market.

^d Time period before duty was put in place.

^e Time period while duty was put in place.

^f Time period after duty was suspended.

to shifting demands brought about by, for example, changes in net farm income. As a result, the absolute user rents could increase from one year to the next even if prices increased. This is evident from table 1, where the absolute user rents are given for selected time periods.

The effect of potash duties on U.S. farmers can be seen by comparing the user rents between 1986 (1) and 1988 (1). Agricultural sales of muriate of potash were 2.499 million metric tons of potassium oxide, which was very close to the agricultural sales of 2.314 million metric tons in 1988 (1). User rents decreased by roughly \$73 million between 1986 (1) and 1988 (1).

In summary, the effects from the U.S. anti-dumping litigation against Canada during the period 1986 (1) to 1988 (1) were:

(1) U.S. Potash Producers	+\$ 12.9 million
(2) Canadian Potash Producers	+\$108.4 million
(3) U.S. Farmers	-\$ 70.4 million
Net Effect	+\$ 50.9 million

From the above, the Canadian firms were the

largest gainers from resolving the potash dispute between the United States and Canada. The losers were U.S. farmers. There was a net gain of roughly \$51 million. This is a conservative estimate since, by comparing other time periods, the Canadian potash producers gained much more than the \$108.4 million shown above. The above numbers show that, in aggregate, the "gainers" gained more from the lawsuit than the "losers" lost. From a U.S. perspective, however, it would have been in the best interest of U.S. farmers to bribe the U.S. potash firms to the tune of roughly \$15 million in order for them not to instigate legal procedures. It follows that, from a U.S. perspective, the "gainers" gained much less than the "losers" lost.

Pricing Strategy of Canadian Firms

In this section, we present results for selected time periods to determine the type of pricing behavior carried out by Canadian firms (table 2).

Table 2. Potash: Effects of Different Price Strategies (selected years)

Time Period and Price Strategy ^a	U.S. User Surplus	Producer Surplus	
		Canada	U.S.
		----- (U.S. \$ thou.) -----	
Prior to duty ^c			
1983 (2)			
Competitive strategy ^b	626,779	0	-18,348
Actual strategy	559,877	52,726	1,477
Dominant strategy	156,695	123,371	171,671
1985 (1)			
Competitive strategy	566,209	0	-18,348
Actual strategy	513,917	40,979	-744
Dominant strategy	141,552	102,501	162,256
1986 (1)			
Competitive strategy	532,311	0	-18,348
Actual strategy	517,223	12,962	-5,386
Dominant strategy	133,078	91,041	156,766
Interim duty ^d			
1987 (2)			
Competitive strategy	265,831	0	13,038
Actual strategy	298,673	-26,388	-901
Dominant strategy	66,458	9,167	136,787
Post duty ^e			
1988 (1)			
Competitive strategy	537,054	0	-18,348
Actual strategy	443,478	71,543	7,526
Dominant strategy	134,263	92,634	157,545

Source: Calculated

Note: Price strategy effects are based on the prices and quantities generated for the competitive, actual, and dominant strategies; elasticity of demand = -0.4

^a (1) refers to the period January-June; and (2) refers to the period July-December.

^b Under the competitive situation, the Saskatchewan firms have zero producer surplus since they set a price equal to marginal cost. The U.S. producers incur a loss if they continue to produce since their marginal cost is greater than the marginal cost of Saskatchewan producers.

^c Time period before duty was put in place

^d Time period while duty was put in place.

^e Time period after duty was suspended.

The rents generated by the actual strategy are compared with the outcomes from two alternative models: (a) Canadian potash producers act competitively; (b) they act as a dominant firm. The theoretical base was developed earlier. The results show that the actual behavior of potash firms is neither competitive nor dominant. The actual behavior seems to deviate significantly from either a competitive or dominant strategy. In comparing the dominant strategy with either the competitive or actual strategies, the rent effect for U.S. producers is much greater than for Canadian producers. That is, a dominant firm strategy increases the rents to U.S. firms to a greater extent than for Canadian producers. The theoretical model supports this conclusion because one is assuming that U.S. firms under a dominant pricing strategy continue to operate at capacity. However, in interpreting these results, the demand elasticity for potash plays a key role. The estimates in the table are based on an inelastic demand schedule for potash. However, the rents would be much lower from a dominant pricing strategy if the price elasticity for potash demand were elastic.

From a U.S. producer perspective, clearly, the dominant firm strategy generates the worst possible outcome. For example, under a competitive strategy, the user rents are at least three times greater than under a dominant firm strategy.

Conclusion

The two U.S. potash firms gained sizable rents from their rent-seeking strategy of lobbying their government to bring countervailing duty action against Canada. But the Canadian producers also gained as the U.S. removed the duties in response to Canada raising its export price. The U.S. antidumping action resulted in a cooperative solution where both the U.S. and Canadian producers gained rents at the expense of the U.S. users.

Our results show that the North American potash producers gained more from the U.S. antidumping action against Canada than what U.S. farmers lost. However, from a U.S. perspective, potash producers gained less than what farmers lost. Usually, in agricultural border disputes, the opposite is the case, in that they involve farm producers who gain by bringing antidumping charges against exporters.

Fortunately for U.S. farmers, even with the

threat of tariffs reappearing—in the event that future dumping might occur—Canadian potash producers are not using a dominant firm pricing strategy in the North American market. If the theoretical dominant firm strategy were followed, the U.S. producers would incur significant additional costs.

[Received February 1990; final revision received August 1990.]

References

- Anderson, D. *The Saskatchewan Potash Industry: Alternative Strategies for Future Development*. Ottawa, Ontario: Economic Council of Canada Pub. No. 264, May 1984.
- Bennett, J. "An Econometric Simulation Model of the Potash Industry." Ph.D. thesis, Tufts University, 1984.
- Bowley, D. "Carlsbad Takes on Canada in Suit." *Carlsbad Current-Argus*, Carlsbad, New Mexico, 10 Feb. 1987.
- Bredahl, M., A. Schmitz, and J. Hillman. "Rent Seeking in International Trade: The Great Tomato War." *Amer. J. Agr. Econ.* 69(1987):1–10.
- Ethier, Wilfred J. "Dumping." *J. Polit. Econ.* 90(1982):487–506.
- Flatters, F., and N. Olewiler. *Dominant Government Firms in an Oligopolistic Industry: The Case of Saskatchewan Potash*. Centre for Resource Studies No. 29, Queen's University, Jan. 1984.
- Fulton, M., and A. Schmitz. "Factors Influencing Output and Investment Decisions in the World Potash Market." A report prepared for the Potash Corporation of Saskatchewan, Saskatoon, Saskatchewan, June 1987.
- Greenaway, N. "Settlement Ends Threat of Penalty for Potash." *The Globe and Mail*, Toronto, Ontario, 8 Jan. 1988.
- Heady, Earl O., and M. H. Yeh. "National and Regional Demand Functions for Fertilizer." *J. Farm Econ.* 41(1959):332–38.
- Hee, Olman. *A Statistical Analysis of U.S. Demand for Phosphate Rock, Potash and Nitrogen*. Washington DC: U.S. Department of Interior, Bureau of Mines Info. Circ. No. 8418, 1969.
- Korol, M. H. "Pricing Behaviour in the Potash Industry: The Dominant Firm, Model Limit Pricing, and Price Discrimination." M.Sc. res. proj. University of Saskatchewan, 1987.
- Lloyd, Peter J. *Anti-dumping Actions and the GATT System*. London: Trade Policy Res. Centre Thames Essay No. 9, 1977.
- Marhatta, Hari. "The Economics of Fertilizer: Alternatives for Avoiding a Shortage." Ph.D. thesis, University of Connecticut, 1976.
- Marud, M. "Potash Producers Still Anger Farmers." *Star-Phoenix*, Saskatoon, Saskatchewan, 19 Jan. 1988.

Olewiler, N. *The Potash Corporation of Saskatchewan: An Assessment of the Creation and Performance of a Crown Corporation*. Ottawa, Ontario: Economic Council of Canada Pub. No. 303, April 1986.

Schmitz, Andrew, Robert S. Firch, and Jimmye S. Hill-

man. "Agricultural Export Dumping: The Case of Mexican Winter Vegetables in the U.S. Market." *Amer. J. Agr. Econ.* 63(1981): 710-19.

U.S. Department of Interior, Bureau of Mines. *Mineral Industry Survey*. Washington DC, 1981-87.

Effects of Technological Change and Institutional Reform on Production Growth in Chinese Agriculture

Shenggen Fan

Recent rapid agricultural production growth in Chinese agriculture could be attributed to an increase in inputs, technological change, and institutional reform. An accounting approach was used to separate the relative contribution of these three factors.

Institutional change, like the introduction of the household production responsibility system, has contributed to past growth in production. However, technological change is crucial to furthering production growth because of the limited potential for significant increase in the use of conventional inputs, in particular land. Continued institutional change must accompany corresponding technological changes.

Key words: agricultural production growth, China, frontier production function, institutional change, regional growth, technological change, total factor productivity.

From 1949 to 1986 agricultural production grew 4% a year in China (Fan). This growth was the most rapid among all the socialist countries (Wong) and even more rapid than growth in most developing countries (Hayami and Ruttan). Contributing to the rapid production growth was a series of technological and institutional changes and rapid increase of modern inputs. Since 1979, efforts have been made to improve incentives and stimulate production by decentralizing authority and responsibility for production decision to family units. Substantial improvement in productive efficiency has resulted.

Using a traditional accounting approach initiated by Solow, Perkins and Yusuf, and Wiens measured the total factor productivity in Chinese agriculture; however, the sources of productivity growth in their studies were not identified. Recently, some studies have measured the effects of institutional change on production and productivity growth. Lin (1987) attributed the

rapid growth in agricultural production from 1980 to 1984 to the household production responsibility system. He found that 20% of productivity growth, or 60% of agricultural production growth, was attributed to the institutional change. However, he ignored the effects of technological change on production and productivity growth. McMillan, Whalley, and Zhu used the accounting approach to capture the effects of reforms in prices and incentive systems on total productivity growth. Their results suggest that 22% of the increase in productivity in China's agriculture between 1978 and 1984 resulted from higher prices and 78% from change in the incentive system. They also ignored the effects of technological change.

The purposes of this study are to develop a new approach to capture the relative contributions of input growth, technological change, and organizational reforms to growth of agricultural production and to apply the approach to the major agricultural production regions of China. During the 1950s, the Chinese government divided the country into six administrative regions. This division is inappropriate for an analysis of agricultural productivity. However, formulating regions on differences in land use is not feasible because of data limitations. Therefore, in this study the country is divided into seven regions that take into account the availability of the agricultural data, the geographical features, and the current social and

At present, Shenggen Fan is jointly a research associate at the International Service for National Agricultural Research, the Hague, Netherlands, and at the Center for International Food and Agricultural Policy, University of Minnesota.

Published as Contribution No. 18172 of the series of the Minnesota Agricultural Experiment Station based on research supported by the station and the Rockefeller Foundation.

The author is grateful to Vernon W. Ruttan for his guidance and insights in the development of this paper; to G. Edward Schuh, Willis Peterson, and Karen Brooks for their comments; and to Sylvia W. Rosen for editorial assistance. He also thanks two anonymous reviewers for their extensive help.

cultural conditions. These regions adhere closely to the administrative division and are as follows: (a) Northeast (NE): Heilongjiang, Liaoning, and Jilin provinces. (b) North (N): municipalities of Beijing and Tianjin; Hebei, Henan, Shandong, Shanxi, Shaanxi, and Gansu provinces. (c) Northwest (NW): autonomous regions of Nei Monggol, Ningxia, Xinjiang, and Tibet; Qinghai province. (d) Central (C): Jiangxi, Hunan, and Hubei provinces. (e) Southeast (SE): Shanghai municipality; Jiangsu, Zhejiang, and Anhui provinces. (f) Southwest (SW): Sichuan, Guizhou, and Yunnan provinces. (g) South (S): Guangxi autonomous region; Fujian and Guangdong provinces.¹

Effects of Input Growth, Technological Change, and Efficiency Improvement

In traditional productivity theory, total production growth consists of movements along the production function (an increase of total inputs) and shifts of the production function (technological change), assuming that the firm is perfectly efficient in production. The growth rate of total factor productivity is the growth rate of total output minus the growth rate of total input; hence, technological change is considered the unique source of productivity growth, and the effects of efficiency improvement on productivity growth are ignored. The assumption of perfect efficiency in production is unrealistic. Differences among firms between realized output and potential output are caused by differences in the capacity to use new technological knowledge and in the motivations of farmers. If this assumption is relaxed, total production growth can be attributed to efficiency improvement as well as to increased inputs and technological change. Different policy inferences may be drawn inasmuch as technological change and efficiency improvement represent fundamentally different sources of growth in production. Therefore, a new approach will be developed to capture all three effects on production growth.

In this study technological change is defined as a shift of the frontier production function. Efficiency improvement is defined as the decrease in the distance between the firm's realized output and its potential output (or frontier). The different sources of production growth are shown in figure 1. At times 1 and 2, the producer faces

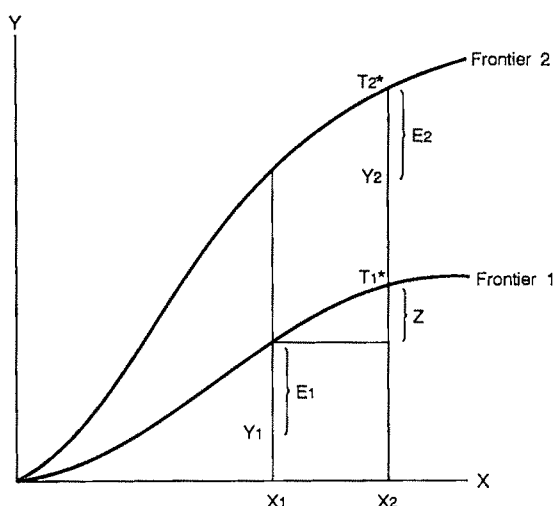


Figure 1. Effects on production growth of input increase, technological change, and efficiency improvement

production frontiers 1 and 2, respectively. If production were perfectly efficient, output would be T_1^* at time 1 and T_2^* at time 2. However, the producer's realized output is Y_1 at time 1 and Y_2 at time 2 owing to production inefficiency. Technological change is measured by the distance between frontier 2 and frontier 1, i.e., $T_2^* - T_1^*$. Inefficiency is measured as the distance between the frontier and the output realized by the producer, i.e., E_1 at time 1 and E_2 at time 2. Hence, the improvement of efficiency over time is the difference between E_1 and E_2 . The contribution of input change is measured as Z . Therefore, the total production growth can be decomposed to three effects: input growth, technological change, and efficiency improvement.

$$Y_2 - Y_1 = Z + (T_2^* - T_1^*) + (E_1 - E_2).$$

Prior to the introduction of the household production responsibility system to Chinese agriculture, production was organized by production teams or state farms. A farmer's income was not closely related to his production effort. After the reform, when producers became responsible for their plots, they worked harder, allocated resources more efficiently, and produced more output with the same input and technology. If only technological change is considered as the source of production and productivity growth, the effects of technological change will be overestimated by ignoring institutional change. Therefore, efficiency improvement is used in this study to capture the effect of insti-

¹ Hainan was not separated from Guangdong province.

tutional change on production and productivity growth.

Frontier Production Function

The frontier production function approach, initiated by Farrell in 1957, has been expanded by various methods of measuring and computing production functions and efficiency (Lovell and Schmidt). The main approaches include pure programming, modified programming, the deterministic statistical frontier, and the stochastic frontier. Pitt and Lee indicated that the programming approach and the deterministic frontier approach do not allow for random shocks in the production process; as a result, a few extreme observations can determine the frontier and exaggerate the maximum possible output. In this study, the stochastic frontier approach is employed to avoid this problem.

Consider the following production function:

$$(1) \quad Y_{it} = f(x_{it}, b)e^{v_{it}}e^{u_{it}}, \text{ or} \\ \ln Y_{it} = \ln f(x_{it}, b) + v_{it} + u_{it}$$

where i denotes the i th firm or region, and t denotes time t ; Y_{it} is output; x_{it} is $1 \times k$ rows of inputs; $f(x_{it}, b)$ is potential output; v_{it} is a stochastic variable representing uncontrolled random shocks in production; and u_{it} is one-sided distribution, $u \leq 0$, which represents technical inefficiency. In (1), $f(x_{it}, b)e^{v_{it}}$ is the stochastic frontier, given that v_{it} consists of random factors outside the firm's control. The nonpositive disturbance u indicates that output must lie on or below the frontier $f(x_{it}, b)e^{v_{it}}$ because $e^{u_{it}}$ has a value between zero and one. It is assumed that for $t \neq t'$, $E(u_{it}u_{it}') = 0$ for all i , and $E(u_{it}u_{it}') = 0$ for all $i \neq j$. In this specification, the firm's inefficiency may change over time by learning from experience. We also assume u is truncated normal with variance σ_u^2 , v is normal with mean zero and variance σ_v^2 , and $E(u_{it}v_{it}') = 0$.

The efficiency for a firm or region i at time t , then, is defined as

$$\frac{Y_{it}}{f(x_{it}, b)e^{v_{it}}}$$

Based on the conditional distribution of u_{it} , given the distribution $v_{it} + u_{it}$, the efficiency of a specific firm or region at a given time can be measured as (Kalirajan and Flinn)

$$(2) \quad E\left\{\exp\left(\frac{u_{it}}{u_{it} + v_{it}}\right)\right\} \\ = \exp\left[-\left(\frac{\sigma_u \sigma_v}{\sigma}\right)\left(\frac{f(\cdot)}{1 - F(\cdot)} - \frac{\varepsilon_{it}}{\sigma} \frac{\sqrt{\lambda}}{1 - \lambda}\right)\right],$$

where $\varepsilon_{it} = v_{it} + u_{it}$, σ is standard error of ε_{it} , $\lambda = \sigma_u^2/\sigma^2$, and $f(\cdot)$ and $F(\cdot)$ are the values of the standard normal density function and standard normal distribution function evaluated at

$$\frac{\varepsilon_{it}\sqrt{\lambda}}{\sigma(1 - \lambda)}.$$

The next step of the specification is to choose an appropriate functional form. Consider a production process that uses n inputs to produce one output represented by the production function

$$(3) \quad Y = f(x_1, \dots, x_n, T),$$

where Y is output, x_i is i th input and T is used to catch technical progress (time trend). The unrestricted translog form can be used to represent production function (3). However, the translog form needs considerable data and has many variables which may lead to multicollinearity problems. Consider a restriction that all inputs are separable from each other, but each input cannot be separated from technical progress:

$$(4) \quad Y = f\{g_1(x_1, T), \dots, g_n(x_n, T)\}.$$

The theoretical background of this form comes from the fact that every input changes over time while the effects among inputs are indirect through time. Then, the following production function form can be used to represent (4):

$$(5) \quad \ln(Y) = a_0 + a_t t + \sum_i a_i \ln(x_i) \\ + \sum_i a_{it} \ln(x_i) \times t + a_{it} t^2.$$

If all inputs and time are considered separable, the production function can be expressed as

$$(6) \quad Y = f\{g_1(x_1), \dots, g_n(x_n), T\}.$$

The Cobb-Douglas production function can be used to represent (6)

$$(7) \quad \ln(Y) = a_0 + \sum_i a_i \ln(x_i) + a_t t.$$

Owing to the serious multicollinearity prob-

lem of the translog form and the constancies of production elasticities in the Cobb-Douglas form, functional form (5) is used for the estimations. The Cobb-Douglas form and average production functions are also estimated for comparison purposes.²

Estimation of Production Functions and Efficiency

Panel data from twenty-nine provinces, municipalities, and autonomous regions in 1965, 1970, 1975, 1976, through 1986 are used in the estimations. Gross agricultural production value serves as the aggregate total output using 1980 constant prices. The subaggregates are (a) crop production, (b) forestry, (c) animal husbandry, (d) sideline industries, and (e) fisheries. Rural industry at all levels (including town, village, and teams) is excluded from agricultural production.³

Labor input in agriculture is measured by the numbers of employed persons at year end.⁴ The sum of sown areas and pasture is used to measure land input because the arable land data are inaccurate. Pasture areas are calculated in sown land area equivalence for output value, i.e., one unit of pasture equals .0124 of a unit of sown land (in 1985).⁵ Chemical fertilizer input is measured by pure nutrients, using the following percentage: 20% for ammonium sulfate, 18.7% for super phosphate, and 40% for potassium sul-

fate.⁶ Machinery input is measured by total horsepower at year end.⁷

Manurial fertilizer, which always has been very important in China, includes animal, human, and crop wastes; green manures; and water plants. In this study, manurial fertilizer is measured from the agricultural population (i.e., human waste) and numbers of domestic animals.⁸ Draft animals are measured at year end in units of heads that are used for agricultural activities and rural transportation. They include water buffaloes, cattle, horses, asses, mules, and camels.⁹ Irrigation input is measured as irrigated areas.¹⁰

The results of production function estimation for the different specifications are shown in table 1. The ordinary least squares technique is used for the average production function estimation and the maximum likelihood technique for the frontier production function. The Cobb-Douglas form is used for regressions 1 and 2. Time trend (T) measures neutral technological change over time. Except for machinery and irrigation, the coefficients of regressions 1 and 2 are very significant considering the crudeness of the data. However, the negative coefficients of draft animals are unrealistic. The sum of production elasticities of traditional inputs (except for draft animals) is more than .75, which implies that traditional inputs still dominate China's agricultural production. Chemical fertilizer input plays an important role in production. The significant and positive time trend coefficient

² The traditional estimation of a production function assumes that every firm is technically efficient, resulting in the average production function, i.e., $Y_{it} = f(x_{it}, b)e^{u_{it}}$, where u_{it} has normal distribution, $N(0, \sigma^2)$.

³ The time series of provincial monetary value of total production (measured in 1980 constant prices) before 1985 is reported in *Collection of Statistical Materials in National Income, 1945-1985*, State Statistical Bureau. The data after 1985 are reported in *China's Statistical Yearbooks*, 1986, 1987, State Statistical Bureau.

⁴ The provincial data of labor before 1980 are calculated from the provincial agricultural population.

$$L_{it} = P_{it} \times \frac{r_{i,80}}{r_{n,80}} \times r_{n,t},$$

where L_{it} denotes i th region's labor input in year t ; P_{it} , i th region's population in year t ; $r_{i,80}$, i th region's ratio of labor to population in year of 1980; $r_{n,80}$, national ratio of labor to population in year 1980. $r_{n,t}$, national ratio of labor to population in year t . The data for agricultural population before 1980 are taken from *National Agricultural Statistical Materials for 30 Years (1949-1979)*, State Statistical Bureau. The data of agricultural labor after 1980 are taken from various issues of China's statistical yearbooks.

⁵ The data for sown areas and pasture are taken from *National Agricultural Statistical Materials for 30 Years (1949-1979)*, State Statistical Bureau.

⁶ The data before 1980 are reported in *National Agricultural Statistical Materials for 30 years (1949-1979)*. The data after 1980 are taken from various issues of China's statistical yearbooks.

⁷ The horsepower of 1965 and 1970 is interpolated based on the numbers of hand tractors and other tractors. The horsepower from 1970 to 1975 is taken from the *National Agricultural Statistical Materials for 30 Years (1949-1979)*. The horsepower after 1980 is taken from various issues of the statistical yearbooks.

⁸ The FAO estimated that one animal (horse unit) produces about 4 tons of manure per year and a person produces .25 ton per year. Manure contains 2.2% pure nutrient, and the manure availability is about 75% of total use. Therefore, manurial resources are estimated as follows:

$$\begin{aligned} \text{Annual manurial resources (tons)} \\ = ((.25 \times \text{rural population} + 4 \\ \times \text{numbers of livestock}) \times 2.2\%) \times 75\%. \end{aligned}$$

The results of this estimation are not significantly different from that of Stone (Tang and Stone).

⁹ The numbers of draft animals before 1980 are taken from the *National Agricultural Statistical Materials for 30 Years (1949-1979)*. The numbers after 1980 are taken from various issues of statistical yearbooks after 1980.

¹⁰ The data of irrigated areas before 1980 are reported in *National Agricultural Statistical Materials for 30 Years (1949-1979)*. Those after 1980 are published in the various issues of statistical yearbooks.

Table 1. Estimates of Production Functions

Regression No:	R1 (Average)	R2 (Frontier)	R3 (Average)	R4 (Frontier)	R5 (Average)	R6 (Frontier)
Constant	-2.81 (10.72) ^a	-2.70 (-11.27)	-2.81 (-5.23)	-3.19 (-6.13)	-2.92 (-6.24)	-2.82 (-6.14)
LABOR	.278 ^{ab} (7.19)	.266* (6.14)	.420* (5.16)	.417* (4.66)	.438* (5.40)	.428* (4.94)
LAND	.356* (7.88)	.379* (9.39)	.243* (2.40)	.331* (3.99)	.246* (2.78)	.261* (3.60)
C. FERT ^c	.235* (8.71)	.236* (9.29)	.140* (2.70)	.089*** (1.66)	.132* (2.57)	.132* (2.61)
MACHINERY	.055** (1.77)	.051** (1.82)	.078*** (1.39)	.123* (2.52)	.075*** (1.35)	.068*** (1.30)
M. FERT ^c	.185* (5.30)	.178* (5.67)	.227* (2.99)	.266* (3.27)	.241* (4.18)	.241* (3.40)
ANIMALS	-.132* (-5.13)	-.133* (-4.94)	.002 (.037)	-.026 (-.301)		
IRRIGATION	.059** (1.81)	.055** (1.66)	.009 (.145)	-.037 (-.537)		
T ^c	.0123* (2.41)	.0125* (2.17)	.0014 (.364)	.0420 (.980)	.0496 (1.28)	.0505*** (1.33)
LABORT ^c			-.0097** (-1.822)	-.0109** (-1.79)	-.0111* (-2.07)	-.0108** (-1.83)
LANDT ^c			-.0024 (-.368)	-.0065 (-1.20)	-.0073 (-1.25)	-.0077*** (-1.64)
C. FERTT ^c			.0068** (1.83)	.0087* (2.41)	.0083* (2.23)	.0081* (2.30)
MACHINERYT ^c			.0080** (1.93)	.0083* (2.08)	.0092* (2.33)	.0098* (2.56)
M. FERTT ^c			-.00006 (-.013)	-.0014 (-.273)	-.0050 (-1.27)	-.0051 (-1.13)
ANIMALST ^c			-.006 (-1.51)	-.0041 (-.725)		
IRRIGATIONT ^c			-.0003 (-.064)	.0006 (.118)		
T ²			.00147* (2.23)	.0013* (2.30)	.0012** (1.80)	.0011*** (1.58)
λ		.822* (2.17)		1.278* (3.23)		.821*** (1.56)
σ		.288* (9.38)		.266* (10.99)		.254* (6.84)
Observations	406	406	406	406	406	406
R ²	.940	.932	.957	.942	.954	.959

^a Numbers in parentheses are *t*-test values.

^b Single asterisk indicates significant at 5% level, double asterisk indicates significant at 10% level, and triple asterisk indicates significant at 20% level.

^c C. FERT is chemical fertilizer; M. FERT, manurial fertilizer; T, Time Trend, $T = 1$ for 1965, $T = 6$ for 1970, ... $T = 22$ for 1986; LABORT, cross term of labor and time trend; LANDT, cross term of land and time trend ...; IRRIGATIONT, cross term of irrigated areas and time trends.

strongly suggests that total factor productivity in Chinese agriculture has increased through neutral technological change.

Functional form (5) is used for regressions 3, 4, 5, and 6. Production elasticity for input i in this production functional form is $\partial \ln Y / \partial \ln x_i = x_i + a_{it}$. Thus, if $a_{it} > 0$, production elasticity of input i is increasing; if $a_{it} < 0$, production elasticity of input i is decreasing.

Regressions 3 and 4 use the same input variables as regressions 1 and 2. In addition, the

cross-term of each input and time trend captures the relative changes of each input in total input over time. The greater significance of the coefficients in regression 4 relative to those in regression 3 implies that the frontier production function used for estimation improved the results. Labor, land, draft animals, and manurial fertilizer play a decreasing role in production, whereas production elasticities of chemical fertilizer and machinery increase over time.

Because the coefficients of draft animals are

negative and the irrigation coefficients are not significant in regressions 1 through 4, these two variables are omitted in regressions 5 and 6. Some effects of draft animals on production are reflected by manurial fertilizer. The improvement in irrigation in China mainly occurs through increased irrigation power rather than an expansion in the size of irrigated areas. Therefore, these omissions do not greatly affect the estimation. Furthermore, these omissions avoid the collinearity among draft animals, manurial fertilizer, and land input. Most of the estimators in regressions 5 and 6 are significant. The omissions of draft animals and irrigation did not cause changes in other coefficients. Again, the frontier estimation is superior to the average estimation.

Table 2 shows that production elasticities (calculated using regression 6) of traditional inputs—land, labor, and manurial fertilizer—are decreasing: labor by 3.6% per year; land, 4.6%; and manurial fertilizer, 3.1%. The annual rates of increase of production elasticities for modern inputs—machinery, 6.5%; chemical fertilizer, 3.9%—are greater than the rates of decrease for traditional inputs.

The results in table 2 can be compared to those of other studies. For example, Ma, Calkins, and Johnson estimated the production elasticities (using 1984 data) for Shuyang county, Jiangsu province. The ranges in value for their elasticities were as follows: labor, .25 to .36; land, .17 to .20; chemical fertilizer, .17 to .23; manurial fertilizer, .08 to .11; and other inputs, .22 to .29. The elasticities vary depending on crops. Wong's estimation of the production functions (using 1960–80 data) for nine socialist countries resulted in the following production elasticities:

labor, .223; land, .143; chemical fertilizer, .177; machinery, .122; and livestock, .233. Comparing those to the production elasticities in table 2, we observe that the elasticities of land and labor in China are greater than those in the socialist countries, indicating that Chinese agriculture uses more traditional inputs than other socialist countries.

The level and variability of technical efficiency for each region are calculated in table 3, using (2) and the results of the frontier production function from regression 6. During the 1960s and 1970s, technical efficiency was about 70%. Efficiency has improved significantly since the institutional change in 1979. The institutional change has three effects: (a) Farmers' incomes and efforts have been linked through improved incentive systems. (b) Farmers may leave agriculture to engage in nonagricultural activities (mainly rural industry), thus improving the land/labor ratio. (c) Farmers may allocate their time and resources to produce high-profit crops, which has improved allocative efficiency and the full use of regional comparative advantages.

It is widely accepted that the introduction of the household production responsibility system enlarged the differences in income among regions (Jiang and Luo). However, there is no evidence that the differences in productive efficiency have increased—the coefficient of variation in productive efficiency has decreased since the reform (see the last column of table 3). The disparity between the production efficiency improvement and income growth among regions suggests that the substantial improvement in production efficiency in poor regions owing to the recent institutional reform did not result in a corresponding increase in income. One reason for this lack of response is the distorted prices in agriculture. Despite the substantial increase in prices in the last ten years, the agricultural product prices still are not reflected by supply and demand. Further reform in prices is needed to give farmers greater incentive to promote further production growth. Another reason is the uneven development of rural industry. The low level of income per capita, especially in the Southwest, is the result of the underdevelopment of rural industry.

Accounting for Total Production Growth

In this part an empirical approach is developed and used to separate the effects on production

Table 2. Production Elasticities for Different Inputs, 1965–1985

	Labor	Land	Chemical Fertilizer	Machinery	Manurial Fertilizer
1965	.417	.253	.140	.078	.235
1970	.363	.215	.181	.127	.210
1975	.309	.176	.221	.176	.185
1976	.298	.168	.229	.186	.180
1977	.287	.161	.237	.195	.174
1978	.276	.153	.246	.205	.169
1979	.265	.145	.254	.215	.164
1980	.254	.138	.262	.225	.159
1981	.244	.130	.270	.234	.154
1982	.233	.122	.278	.244	.149
1983	.222	.114	.286	.254	.144
1984	.211	.107	.294	.264	.139
1985	.200	.099	.303	.274	.134

Table 3. Level and Variability in Technical Efficiency of Seven Regions for Selected Years

Year \ Region	NE	N	NW	C	SE	SW	S	National Average	C.V. ^a
1965	.868	.433	.698	.728	.679	.681	.644	.646	.191
1970	.853	.561	.844	.844	.847	.731	.846	.772	.138
1975	.887	.581	.808	.881	.866	.652	.812	.761	.127
Average 65-79	.892	.574	.758	.850	.817	.713	.789	.737	
Rank	1	7	5	2	3	6	4		
C.V. 65-79	.033	.117	.103	.069	.084	.061	.087	.132	
Rank	7	1	2	5	4	6	3		
1980	.917	.625	.692	.826	.802	.781	.756	.753	.122
1981	.911	.630	.774	.858	.851	.791	.758	.768	.114
1982	.911	.645	.777	.885	.863	.851	.810	.788	.109
1983	.939	.681	.751	.863	.847	.858	.795	.791	.103
1984	.934	.726	.799	.908	.900	.894	.831	.831	.070
1985	.891	.725	.829	.909	.906	.891	.870	.843	.076
Δ 70s-85 ^b	.001	.151	.071	.059	.089	.178	.081	.106	
Rank	7	2	5	6	3	1	4		
Δ 65-85	.023	.292	.131	.181	.227	.210	.226	.197	
Rank	7	1	6	5	2	4	3		
Average 65-85	.898	.616	.766	.863	.844	.771	.807	.772	
Rank	1	7	6	2	3	5	4		
C. V. 65-85	.033	.123	.081	.056	.073	.105	.081	.130	
Rank	7	1	3	6	5	2	4		

^a C. V. is coefficient of variation^b Δ 70s-85 indicates the absolute improvement of technical efficiency between 1965-79 average and 1985.

growth of an increase in inputs, technological change, and institutional reform. Using functional form (5), the production function can be expressed as

(8)

$$\ln Y(t) = a_o + \sum_i a_i \ln x_i(t) + \sum_i a_{it} (\ln x_i(t)) \times t \\ + a_t t + a_{tt} t^2 + \ln(e^{u(t)}) + v(t)$$

$$(9) = \ln A_o(t) + \sum_i a_i(t) \ln x_i(t) + \ln E(t),$$

where $\ln A_o(t) = a_o + a_t t + a_{tt} t^2 + v(t)$, $a_i(t) = a_i + a_{it} t$, and $E(t) = e^{u(t)}$.

Taking the first derivative of (9) with respect to time t , the growth rate of total production can be accounted for as

$$(10) \quad \partial \ln Y(t) / \partial t = \partial \ln A_o(t) / \partial t + \sum_i a_i(t) \\ \times \partial \ln x_i(t) / \partial t + \sum_i \ln x_i(t) \\ \times \partial a_i(t) / \partial t + \partial \ln E(t) / \partial t.$$

The first term in (10) measures neutral tech-

nological change. The second term captures the effect of input change on production growth; it is the sum of growth rates in inputs weighted by the relevant production elasticities. The third term measures the effects of biased technological change on production growth; if it is positive, output has increased through biased technological change (using abundant resources to substitute for scarce resources). The last term reflects the effect of institutional change (or efficiency improvement) on production growth.

Using (10), the accounting for the sources of total production growth is presented in table 4. Neutral and biased technological change are considered as total technological change in the accounting and treated as the residual. For the whole country, total production growth rate was 5.04% per year from 1965 to 1985; 57.7% of the growth is explained by increased use of total input and 42.3% by growth in total factor productivity. About 63% of productivity change is attributed to institutional change (or efficiency improvement) and about 37%, to technological change. The increase of labor still explains about 7.7% of total production growth. The change of land input had the least effect because acreage

Table 4. Accounting for Growth of Total Agricultural Production in Terms of Annual Growth Rates, 1965–1985

	NE	N	NW	C	SE	SW	S	National
Total Production Growth	5.09	5.88	3.70	4.40	5.50	4.40	4.50	5.04
	(100)	(100)	(100)	(100)	(100)	(100)	(100)	(100)
Total Input Growth	3.10	3.10	2.72	2.71	2.80	3.66	2.55	2.91
	(60.9)	(52.7)	(73.5)	(61.6)	(50.9)	(83.2)	(56.7)	(57.7)
Labor	.23	.24	.45	.43	.25	.67	.49	.39
	(4.5)	(4.1)	(12.2)	(9.8)	(4.5)	(15.2)	(10.9)	(7.7)
Land	.04	-.05	-.07	-.01	.06	.11	0	.002
	(.8)	(-.9)	(-1.9)	(-.2)	(1.1)	(2.5)	(0)	(.04)
C. Fert.	1.73	1.61	1.51	1.22	1.29	1.45	.79	1.32
	(34.0)	(27.4)	(40.8)	(27.7)	(23.5)	(33.0)	(17.3)	(26.2)
M. Fert.	.20	.35	.18	.13	.04	.36	.31	.25
	(3.9)	(6.0)	(4.9)	(3.0)	(.73)	(8.2)	(6.9)	(5.0)
Machinery	.90	.95	.65	.94	1.16	1.07	.96	.95
	(17.7)	(16.2)	(17.6)	(21.4)	(21.1)	(24.3)	(21.3)	(18.8)
Total Productivity Growth	1.99	2.78	.98	1.69	2.70	.74	1.95	2.13
	(39.1)	(47.3)	(26.5)	(38.4)	(49.1)	(16.8)	(43.3)	(42.3)
Institutional Change	.13	2.61	.86	1.11	1.45	.82	1.52	1.34
	(2.5)	(44.4)	(23.2)	(25.2)	(26.4)	(18.6)	(33.8)	(26.6)
Technological Change	1.86	.17	.12	.58	1.25	-.08	.43	.79
	(36.5)	(2.9)	(3.2)	(13.2)	(22.7)	(-1.8)	(9.6)	(15.7)

Note: (10) is employed for the accounting.

used for agriculture remained nearly constant. Among all inputs, increased chemical fertilizer input contributed most significantly to production growth (26.2%), while manurial fertilizer explained 5% of total production growth. The increase in machinery use is the second most important factor in total production increase.

The differences in sources of production growth among regions are substantial because of differences in the resource endowments and total factor productivity growth. Growth in total agricultural production varied from 3.70% in the Northwest to 5.88% in the North region. The contribution of total input growth to production growth varies from 50.9% in Southeast to 83.2% in Southwest. The differences in modern input (chemical fertilizer and machinery) growth explains most of the differences in total input growth. Among modern inputs, chemical fertilizer has the largest effects. The differences in traditional input growth are small.

The differences of the effects of institutional change on production growth explain the largest share of the differences in total production growth, ranging from 2.5% in Northeast to 44.4% in North. The contribution of technological

change to production growth also has varied substantially among regions. Total factor productivity growth in the Northeast is mainly explained by technological change. Technological change contributed more than 45% of the total factor productivity in the Southeast. However, technological change in the North, Northwest, and Southwest contributed little to total factor productivity and total production growth.

Concluding Comments

The major findings of this study are summarized as follows: The estimates of the frontier production functions indicate that traditional inputs are still important to China's agriculture. However, the importance of the traditional inputs of land, labor, and manurial fertilizer is decreasing rapidly. In contrast, the coefficients of modern inputs, e.g., chemical fertilizer and machinery inputs, were small in 1965 but have since increased rapidly. By 1985, the modern inputs were as important as the traditional inputs.

Efficiency measurements indicate that the household production responsibility system has

contributed significantly to production growth. However, the regional differences in performance are large. In general, land-scarce regions gained more from the reform.

The accounting for production growth showed that a significant share of total production growth still can be attributed to increases in traditional inputs. Among all inputs, increased chemical fertilizer use was the most important source of production growth. Increased machinery input ranked second in importance. Total input growth explained 57.7% of total production growth. The residual, the proxy for technological change and efficiency improvement, accounted for 42.3% of total production growth. Institutional change has had greater effects on productivity and production growth than has technological change.

These findings have important policy implications in promoting further production growth and smoothing regional inequalities. China's population reached 1065.29 million in 1987. The population growth rate from 1949 to 1987 was 1.84%, although it declined to 1.29% in last decade. Further decreases in population growth will not be easy in the next decade because the base population is large and those born in the 1960s are entering reproductive age. Thus, the demand for food will continue to grow even apart from income effects. The demand for cash crops is increasing with the development of industrialization. How to meet the future demand for rapid increases in food and in industrial materials is an urgent problem.

Increased Input Use

The quickest solution for China is to increase the use of inputs, such as land, labor, chemical fertilizer, machinery, and others. However, the total land input is likely to decline in the future (Sun). Without an increase in land areas, an increase in labor will have only a limited effect on total production. Increased use of modern inputs, especially chemical fertilizer, likely has the greatest potential for increasing total production. Although fertilizer input per unit of land in China is higher than in most developing countries, the output increase from greater fertilizer use is still potentially large in some regions (see table 4), particularly in the Northeast, Northwest, North, and Southwest. Increased machinery input will have little effect on production unless it increases land productivity. Thus, a top priority in mechanization involves increased land productivity (e.g., mechanization of irrigation).

Technological Change

The results of this study indicate that technological change accounts for 15.7% of total agricultural production growth in China. Compared to other countries, this proportion is very small. In Japan, from 1960 to 1980, 47.4% of total production growth stemmed from technological change, and technological change accounted for 84.2% of the growth in U.S. total output (Hayami and Ruttan).¹¹ Underinvestment in agriculture may explain the slow technological change in China. In 1985, the agricultural sector produced 28.1% of total national output and 41.1% of national income, although the agricultural investment was only 3.4% of total investment (*China's Statistical Yearbook*, 1986). The underinvestment in agriculture has resulted in poor rural infrastructure and insufficient agricultural research. An increase in agricultural investment, especially in research and development, is needed to stimulate technological change.

Institutional Change

Recent institutional changes have improved agricultural production efficiency greatly; 26.6% of production growth has been contributed by institutional change. The new strategy should focus on greater regional specialization, based on comparative advantages. The self-sufficiency policy both at the national level and local levels should be discarded. Crops should be grown where soil and climate provide the most favorable conditions. Although rural labor has more opportunities to work outside agriculture, labor immobility will become a major source of inefficiency. The pattern of land holdings (in terms of size distribution of farm), land tenure and other contractual arrangements in agriculture should be adjusted appropriately to gain more efficiency. The recent introduction of factor and product markets in agriculture has contributed to more efficient allocation of resources; however, instability of input and output prices and the insufficient supplies of modern inputs will continue constrain agricultural production.

¹¹ See table 7.2, Hayami and Ruttan (1985 ed.). Total output growth is 1.9% a year in Japan from 1960 to 1980; total productivity growth (the contribution of technical change to output growth), .9%. Thus, the relative contribution of technical change to total output growth is 47.4%. Using the same calculation, the relative contribution of technical change to total output growth is 84.2% in the United States.

Smoothing Regional Inequalities

Differential growth rates in agricultural development among regions of a country represent a persistent challenge to policy makers. Smoothing the differences in technological and institutional changes among regions is needed to reduce the differences in production and income growth. A well-integrated and extensive physical infrastructure and a strong regional agricultural research capacity adapted to the needs of the regional agricultural economy are important in contributing to development of new comparative advantages in technology in the regions disadvantaged by resource endowments and to stimulation of more even rates of technological change across regions.

New agricultural policies and institutional changes should create more geographically even growth in agricultural production and income. For example, crop prices should be raised in order to narrow the income differences between the regions with an advantage of crop production but with a disadvantage of rural industry, and the regions with well-developed rural industries.

[Received August 1989; final revision received May 1990.]

References

- Agricultural Yearbook Editing Committee. *China Agricultural Yearbook*. Beijing: Agricultural Publishing House, 1980, 1981, 1982, 1983, 1984, 1985, 1986, and 1987, various issues.
- Fan, Shenggen. "Regional Productivity Growth in China's Agriculture." Ph.D. thesis, University of Minnesota, 1989.
- FAO. *China: Recycling of Organic Wastes in Agriculture*. Rome: FAO Soils Bull. No. 40, 1977.
- Farrell, M. J. "The Measurement of Productive Efficiency." *J. Royal Statist. Soc., Series A* 120(1957):253-81.
- Hayami, Yujiro, and Vernon W. Ruttan. *Agricultural Development: An International Perspective*. Baltimore MD: Johns Hopkins University Press, 1971, 1985.
- Jiang, Jinyong, and Xiaopeng Luo. "Changes in the Income of Chinese Peasants since 1978." *China's Rural Development Miracle, with International Comparison*, ed. John W. Longworth. University of Queensland, Australia, 1989.
- Kalirajan, K. P., and J. C. Flinn. "The Measurement of Farm Specific Technical Efficiency." *Pakistan J. Appl. Econ.* 2(1983):167-80.
- Lin, Justin Yifu. "Household Farm, Cooperative Farm, and Efficiency: Evidence from Rural Decollectivization in China." Economic Growth Center Work Pap. No. 533, Yale University, 1987.
- . "The Household Responsibility System in China's Agricultural Reform: A Theoretical and Empirical Study." *Econ. Develop. and Cultur. Change* 36(1988):200-224.
- Lovell, C.A.K., and P. Schmidt. "A Comparison of Alternative Approaches to the Measurement of Productive Efficiency." *Application of Modern Production Theory: Efficiency and Productivity*, ed. Ali Dogramaci and Rolf Färe. Boston: Kluwer Academic Publishers, 1988.
- Ma, Nathaniel J., Peter H. Calkins, and Stanley R. Johnson. "The Household Responsibility System: Technical and Allocative Efficiency vs. Equity." *China's Rural Development Miracle, with International Comparison*, ed. John W. Longworth. University of Queensland, Australia, 1989.
- McMillan, John, John Whalley, and Lijing Zhu. "The Impact of China's Economic Reforms on Agricultural Productivity Growth." *J. Polit. Econ.* 97(1989):781-807.
- Perkins, Dwight, and Shahid Yusuf. *Rural Development in China*. Baltimore MD: Johns Hopkins University Press, 1984.
- Pitt, M. M., and L. F. Lee. "The Measurement and Sources of Technical Inefficiency in the Indonesian Weaving Industry." *J. Develop. Econ.* 9(1981):43-64.
- Solow, Robert. "Technical Change and the Aggregate Production Function." *Rev. Econ. and Statist.* 39 (1957) 312-22.
- State Statistical Bureau. *China Statistical Yearbook*. Beijing, China, various issues. Statistical Press, 1981, 1982, 1983, 1984, 1985, 1986, and 1987.
- . *Collection of Statistical Materials in National Income, 1945-1985*. Beijing: China Statistical Press, 1987.
- . *National Agricultural Statistical Materials for 30 Years (1949-1979)*. Unpublished data. Beijing, 1980.
- Sun, Han. "9.8 Million Square Kilometers of Territory." *Feeding a Billion*, ed. Wittwer Sylvan. East Lansing Michigan State University Press, 1987.
- Tang, Anthony. *An Analytical and Empirical Investigation of Agriculture in Mainland China, 1952-1980*. Taipei: Chung-Hua Institution for Economic Research, 1984.
- Tang, Anthony, and Bruce Stone. *Food Production in PRC*. Washington DC: International Food Policy Research Institute, 1980.
- Wiens, B. Thomas. "Technical Change." *The Chinese Agricultural Economy*, ed. Randolph Barker, Radha Sinha and Beth Rose. Boulder CO: Westview Press, 1982.
- Wong, Lung-Fai. *Agricultural Productivity in the Socialist Countries*. Boulder CO: Westview Press, 1986.

Farmer Behavior Under Risk of Failure

William E. Foster and Gordon C. Rausser

We analyze input decisions under risk of farm failure. Inputs with immediate cash outlays have greater effective than observed prices because their cost increases the probability of failure, and their optimal marginal products are higher than observed prices would warrant under strict profit maximizing without failure risk. An algebraic example illustrates the market equilibrium effects of failure risk. We apply the model to an analysis of Illinois corn production (1971–79). Results indicate that larger farms deviate less than smaller farms from strict profit maximization. Over the period studied, farmers moved further from setting marginal products equal to observed prices.

Key words: corn production, farm failure, input decisions, risk.

This article addresses input decisions under risk of farm failure. Without risk of failure, the farmer would first maximize his expected utility as if he relied only on farm revenue (including concurrent off-farm employment, if applicable) and then compare this level of farm-derived utility with the utility available elsewhere (less moving expenses, etc.). If the farm-derived expected utility exceeds nonfarm utility, the farmer remains a farmer. With risk of failure, however, if the farmer chooses to continue another season, the ensuing revenues may be so low as to preclude future decisions to stay in agriculture (i.e., failure). The critical level of revenue below which the farm fails may be a function of household-maintenance expenditures, farm debt, and so forth. This level is likely to rise with farm size, but at a decreasing rate. Therefore, smaller scale, commercial farms are more likely to be at risk than larger farms.

The discrete difference between on-farm and off-farm returns leads to asset fixity (for example, see Johnson and Quance, Johnson and Pasour). A farmer's human capital fixity, in particular, presents important complications to conventional economic analysis of agricultural supply, which relies on strict maximization of expected utility (or of expected profit) derived solely from farm production. If the difference

between on-farm and off-farm utility is large, the farmer will take actions to avoid leaving agriculture; these actions may lead to seemingly inefficient farm management decisions. Robison, Barry, and Burghardt, for example, explore the use of credit as a means of forestalling costs associated with liquidating fixed assets to meet cash flow obligations. Their model implies that a farmer has a greater incentive to acquire debt as bankruptcy grows more likely, consistent with the "go-for-broke" behavior seen in highly stressed borrowers.

Also important is the influence of the risk of failure on farm input decisions. The purpose of this paper is to draw out the implications of failure risk on farm production decisions, specifically on the differing incentives to use cash-draining inputs in contrast to those inputs that do not require immediate cash outflow. The second section presents a model of input use where there is a difference between on-farm and off-farm returns to a farmer's human capital and shows that, under failure risk, the use of non-cash-draining inputs may increase with a decrease in output price. The third section presents a specific partial equilibrium model to illustrate the concepts developed in the second and to demonstrate the possible effects of failure risk on equilibrium output price. The fourth section turns to the practical application of the ideas developed in this study and to the role commodity programs play in reducing the effect of costly farm failure. The empirical analysis is of corn production in Illinois, using the model of how the threat of failure affects input choices. The

William E. Foster is an assistant professor, Department of Economics and Business, North Carolina State University; Gordon C. Rausser is Robert Gordon Sproul Distinguished Professor, Department of Agricultural and Resource Economics, University of California, Berkeley.

study focuses on the effects of farm size and other structural variables on farmer divergence from strict profit maximization.

Farmer Behavior Under Risk of Forced Adjustment

Two major sources have been hypothesized regarding the fixity of physical assets. The first involves the lumpy nature of the production process. Galbraith and Black, for example, hypothesize that large fixed costs associated with reorganizing the farm operation make short-run adjustment unprofitable. Changes in the economic environment would induce changes in productive capacity only if they were of sufficiently high magnitude and of sufficiently long duration. The second source of fixity involves the difference between the on-farm value of assets and their alternative, off-farm value (Johnson and Quance). Low salvage value may reflect transportation costs, specificity of capital to the farm operation, and limited, asymmetric information regarding quality of the item (Akerlof 1970, 1980). Whatever the cause of low salvage values, once acquired, productive capacity changes only with discrete (perhaps large) changes in the on-farm-use value.

A farmer's human capital is subject to a similar fixity. Although a farmer's labor may be divisible between on-farm and off-farm employment, his human capital specific to farm management is lumpy. Or from another perspective, one may view the "salvage" value of human capital (the opportunity cost of being a farm manager) as less than the on-farm value (the opportunity cost of not being a farm manager). The discrepancy between a farmer's salvage value and his on-farm value may reflect the specificity of the capital and/or the personal, or psychic, premium on earning an income from owning and operating a farm.¹ In dollars (accounting for moving costs, etc.), a farmer may seem to have a high salvage value, but in utility, derived from these nonfarm dollars, the salvage value is low.

A notable distinction between a farmer's human capital and other factors that may be sub-

ject to fixity is that certain minimal expenditures are necessary in order to stay in farming. The farm family must eat, clothe itself, and otherwise live happily alongside its neighbors. In addition, the decision to farm may incur other fixed commitments: minimal debt service, minimal use of certain publicly provided goods (e.g., water), insurance, and so forth.

The adjustment of a farmer's human capital out of agriculture is not always an active decision. Forced adjustment occurs when farm production does not cover the minimal, necessary costs. Without this risk of forced adjustment, or failure, the farmer would simply compare the expected utility of farming with the utility of leaving farming and make the optimal discrete decision to continue. Instead, the risk of failure leads to seemingly inefficient production decisions.

Input response to risk may entail anything from too quickly depleting soil quality, to placing greater stress on farm labor at the expense of leisure.² Inputs that are purchased prior to the production process are underutilized in the sense that their marginal products are higher than what would be optimal under strict maximization. Those that can be utilized without immediate cash expenditure ("mined" or "borrowed off"), such as soil quality and farm household labor, are overutilized. Commodity production may be increased or decreased, relative to the case where there is no difference between on-farm and off-farm utility, depending on the degree of complementarity of factors.

An Algebraic Model

A model of a farmer's behavior under risk of failure is inherently intertemporal. The farmer must trade off the amount of utility he gets in any year from farming with the probability of failure.³ Intertemporal models may lead to intractable complications; therefore, the following mathematical model makes certain simplifying assumptions. Specifically, the farm manager faces a discrete difference between on-farm and off-

¹ One reviewer points out the possibility that the farmer, having failed, may work for other farmers, thereby employing (at least some of) his farm-specific skills. These employed skills, however, would likely be strictly related to farm labor rather than management of inputs (including labor). Nevertheless, the opportunity cost of the farmer's management of a farm may be the value of employment by other farmers rather than employment out of agriculture altogether. Furthermore, the utility of this farm labor may be much less than that derived from farm management.

² For example, Thompson, Gwynn, and Sharp report the survey of married farm women in Yolo County, California, and remark: "The increased participation by women on smaller farms was found to result from the need for the entire family to use its total resource for survival rather than to a greater opportunity for women to participate on small farms."

³ The models presented in this and the following section are similar to those discussed by Just and Zilberman. The basic model offers an explanation of farmer behavior that resembles behavior arising from a safety-first objective function (Pyle and Turnovsky).

farm returns; random events are independent over time; the alternative utility that the farmer receives off the farm, if he fails, is a constant value; and, once failed, the farmer leaves farm management forever. These assumptions produce the following results. Farm decision rules as functions of prices are constant over time; the expected farm-derived utility in any period is constant; the cost of moving out of agriculture is the difference (also constant) between optimal yearly expected farm-derived utility and off-farm utility. Furthermore, the farmer's objective function can be written in terms of yearly expected farm-derived utility, off-farm utility, and the probability of failure.

A farmer yearly produces a commodity, the per-acre amount of which is denoted by y , by combining two factors of production: (a) x are those which must be yearly purchased out of cash revenues prior to realization of actual production and price received, and (b) k are those which may be utilized in the year but paid for in the indeterminate future. Examples of the factor x include hired labor, fertilizer, etc. Examples of input k include land quality, farm household labor, owned machinery, etc. The per-acre production function in a given year is given by

$$(1) \quad y = y(x, k, \varepsilon),$$

where ε represents some random effect on output y , such as weather. The number of acres produced is given by A , which is a constant. The farmer faces every year an unknown price p , a random variable ($p > 0$, $E[p] = \mu$), with some time-invariant probability density function given by $g(p)$. The constant per-unit cost of the factor x is given by w , and of factor k is given by i .

Failure is defined as inability to cover the minimal expenditures in a year necessary to farm, f ; that is, failure is defined by

$$(2) \quad A[py - wx] < f.$$

If the farmer fails, he leaves farming and earns some sure utility level I each year thereafter. For ease of presentation, we assume that production is certain and price is the only random variable. This particular assumption is relaxed in the next section.

The yearly probability of failure, π , is given by $\pi(x, k) = \int_0^{p_d} g(p) dp$, where $p_d y - wx = f/A$. The utility from farming in any year is given by

$$(3) \quad U(p, x, k) = U[A(py - wx - ik)],$$

and the expected utility from farming is simply

Table 1. Future Utility Levels Conditioned on Farm Failure

Discount Factor	Probability of Income Stream				
	π	$\pi(1 - \pi)$	$\pi(1 - \pi)^2$	$\pi(1 - \pi)^3$	
1	U_F	U_N	U_N	U_N	...
β	I	U_F	U_N	U_N	...
β^2	I	I	U_F	U_N	...
β^3	I	I	I	U_F	...
β^4	I	I	I	I	...
...
...

Note: $U_F = E[U | p < p_d]$, and $U_N = E[U | p \geq p_d]$.

$$(4) \quad \bar{U}(x, k) = E[U(p, x, k)].$$

Finally, the constant discount rate for future expected utility (either from farming or not farming) is given by β . Define \bar{U}_N as the expected value of utility from farm income conditional on not failing, \bar{U}_F as the expected value conditional on failure; that is, $\bar{U} = (1 - \pi)\bar{U}_N + \pi\bar{U}_F$.

Table 1 presents possible future events that the farmer must consider if he chooses farming. After summing over all possible streams of utilities,⁴ one may represent the farmer's objective as

$$(5) \quad \bar{V}(x, k) = \frac{\pi \frac{\beta}{1 - \beta} I + \bar{U}(x, k)}{1 - \beta[1 - \pi(x, k)]};$$

The term $I\beta/(1 - \beta)$ is the expected earnings from agricultural work, after failure. (Note the discount factor on this off-farm income is $\beta/(1 - \beta)$, not $1/(1 - \beta)$, because the farmer starts earning I the year following failure.) The term

⁴ Sum along the central diagonal of table 1, taking the product of the expected utility in any year, the probability of obtaining that expected utility, and the appropriate discount factor:

$$\begin{aligned} \bar{U}_F \pi + \bar{U}_F \pi(1 - \pi) + \dots &= \bar{U}_F \sum_{i=0}^{\infty} \beta^i (1 - \pi)^i \\ &= \bar{U}_F \pi / [1 - \beta(1 - \pi)]. \end{aligned}$$

Sum along a representative diagonal j steps to the left of and above the central diagonal:

$$\begin{aligned} \bar{U}_N \pi(1 - \pi)^j \sum_{i=0}^{\infty} \beta^i (1 - \pi)^i \\ &= \bar{U}_N \pi(1 - \pi)^j / [1 - \beta(1 - \pi)], \quad j = 1, 2, \dots \end{aligned}$$

Sum along a representative diagonal j steps to the right of and below the central diagonal:

$$I \beta^j \pi \sum_{i=0}^{\infty} \beta^i (1 - \pi)^i = I \beta^j \pi / [1 - \beta(1 - \pi)], \quad j = 1, 2, \dots$$

Sum all the diagonal summations to obtain expression (6).

$[1 - \beta(1 - \pi)]$ represents the discount rate of future incomes. As the probability of failure increases, the farmer would tend to discount the future more heavily.

A more compact representation of the farmers criterion function is derived by subtracting from \bar{V} in expression (6) the income stream if the farmer left agriculture before the first year:

$$(6) \quad V(x, k) = \bar{V}(x, k) - \frac{1}{1 - \beta} I \\ = \frac{\bar{U}(x, k) - I}{1 - \beta[1 - \pi(x, k)]}.$$

The value $V(\cdot)$ represents the rent, or surplus, from being a farm owner. The numerator, $\bar{U} - I$, represents the farmer's expected surplus in any given year. The value $I/(1 - \beta)$ is the salvage value of a farmer's human capital. When $V(x, k) > 0$, a farmer's human capital is fixed for limited changes in \bar{U} and I .

Let k^* and x^* satisfy the first-order conditions of maximizing the objective function given by expression (7):⁵

$$(7a) \quad \frac{\partial y}{\partial k} = \frac{iE[U']/E[U'p]}{1 + \theta}$$

$$(7b) \quad \frac{\partial y}{\partial x} = wE[U']/E[U'p] \frac{1 + \theta E[U'p]/E[U']p_d}{1 + \theta}$$

$$(7c) \quad \theta = V\beta g(p_d)p_d/[AyE(U'p)].$$

Here, marginal utility is given by U' . The parameter θ measures the influence of human capital fixity (measured by V) on farm production decisions. Expression (8c) makes explicit that the influence of human capital fixity on production depends on several elements: the discount rate, the probability distribution of price, farm size, and the correlation of marginal utility and price. If fixity has no influence on production decisions, then $\theta = 0$; and as θ grows, production decisions deviate from strict maximization.

Compare these first-order conditions to that of strict maximization of $E[U]$:

$$\frac{\partial y}{\partial k} = iE[U']/E[U'p]$$

$$\frac{\partial y}{\partial x} = wE[U']/E[U'p].$$

Because $\theta > 0$, the farmer appears to overutilize k based on conventional marginal conditions. Further, it is reasonable to suppose that the critical price defining failure, p_d , is small relative to expected price. If p_d is such that $E[U'p] > E[U']p_d$, then the farmer appears to underutilize x .⁶

Now suppose the rest of the economy improves relative to the agricultural sector. As the costs of failure ($\bar{U} - I$) grows insignificant, then $\theta \rightarrow 0$, and the farmer behaves as if he were maximizing the utility derived solely from farming (\bar{U}). We may interpret $g(p_d)$ as a measure of the degree to which a farmer can marginally influence the probability of failure via production decisions. For example, if random price were associated with a familiar bell-shaped probability density function, then farmer input decisions decreasing p_d by a unit at low levels would have less effect on the probability of failure than if p_d were decreased by a unit near the mean price. Suppose the probability distribution of commodity price changes such that $g(p_d) \rightarrow 0$ (that is, suppose the bell-shaped density of price shifts rightward with an increase the mean price), then farm decisions move toward production efficiency. Regardless of potential cost of failure (i.e., $\bar{U} - I$), if there is no influence at the margin, $g(p_d) = 0$, then again the farmer acts as if he were maximizing farm-derived utility. One point to be emphasized from this discussion of expression (8) is that in this model human cap-

⁵ Equations (8a), (8b), and (8c) derive from the following first-order conditions for the maximization of the objective function (7):

$$\frac{\partial V}{\partial k} = D^{-1} \frac{\partial y}{\partial k} A[E(U'p) - E(U')] + VD^{-1} g(p_d) \beta p_d \frac{\partial y}{\partial k} y^{-1} = 0, \\ \frac{\partial V}{\partial x} = D^{-1} \frac{\partial y}{\partial x} A[E(U'p) - E(U')] + VD^{-1} g(p_d) \beta p_d \frac{\partial y}{\partial x} y^{-1} \\ - VD^{-1} g(p_d) \beta w y^{-1} = 0,$$

where $D = 1 - \beta(1 - \pi)$.

⁶ The specification of survival—covering minimal expenses—is stylized in one important way. Minimal expenses, f , is a constant, and specifically not an increasing function of three items: farm size, input use, or past failure to cover all expenses (i.e., some portion of past ik). Although relaxing the current simplification would complicate the analysis, to make f a function of the first two items would not alter the basic implications of the model as long as f increases over these items at a decreasing rate. The third item is conceptually more important. If the farmer delays payment on some portion of the cost of inputs k , one expects that portion to contribute to higher minimal expenditures in the future (or perhaps decreased output) and thus a higher future probability of failure for all levels of inputs. This would tend to blur the distinction between x and k in the active avoidance of failure. Nevertheless, as long as the farmer does not have to pay a price of delaying payment that exceeds his discount rate, and at failure time all delayed payments outstanding are forgotten, then k would be a more attractive input than x using the simple criterion of minimizing the probability of failure.

ital fixity ($V > 0$) is necessary but not sufficient to cause "inefficient" production decisions.

With the general representations of $V(\cdot)$ and $g(\cdot)$, one cannot immediately determine the effect of changing farm size on the optimal choices of factors. Although p_d decreases with an increase in A , which decreases θ , $g(p_d)$ may decrease or increase. In addition, V increases with an increase in farm size, positively affecting θ . If, however, p_d is small relative to μ , then $g(p_d)$ likely decreases with farm size (as would be the case if p were normally distributed). The effect of decreasing f , and thus decreasing the probability of failure, would bring k^* and x^* into line with productive efficiency.

To be more specific, suppose the farmer is risk neutral; that is, $E(U')/E(U'p) = 1/\mu$. The first-order conditions may then be written as

$$(8a) \quad \frac{\partial y}{\partial k} = \frac{i}{\mu} / (1 + \theta)$$

$$(8b) \quad \frac{\partial y}{\partial x} = \frac{w}{\theta} \left(1 + \theta \frac{\mu}{p_d} \right) / (1 + \theta)$$

$$(8c) \quad \theta = V\beta g(p_d)p_d / (A\mu y).$$

As noted above, relative to $\theta = 0$, the marginal product of k is set lower and the marginal product of x is set higher. (This is true given $p_d < \mu$.) What happens to the actual levels of k and x depends on the degree of substitutability. For example, suppose k and x represent single factors. Let the parameter θ begin from a point where fixity has no influence on production, either because the cost of failure is zero ($V = 0$), or because the farmer has no influence at the margin over the probability of failure ($g(p_d) = 0$). The following comparative statics show the effect on factor decisions due to an increase in θ , indicating an increase in the influence of fixity. Differentiating (9a) and (9b) with respect to k , x , and θ yields:

$$\frac{\partial k}{\partial \theta} = \frac{-1}{\Delta} \left[y_{xx} \frac{i}{\mu} - y_{xk} \frac{w}{\mu} \left(1 - \frac{\mu}{p_d} \right) \right]$$

$$\frac{\partial x}{\partial \theta} = \frac{-1}{\Delta} \left[y_{kk} \frac{w}{\mu} \left(1 - \frac{\mu}{p_d} \right) - y_{xk} \frac{i}{\mu} \right],$$

where $\Delta > 0$ is the determinant of the matrix of second partial derivatives of y with respect to x and k , which is assumed negative definite. Not surprisingly these effects are of ambiguous sign. However, if the two inputs are substitutes (i.e., y_{xk}), then the effect of increasing the mea-

sure of fixity, V , or of increasing at the margin the farmer's influence on the probability of failure, $g(p_d)$, is to decrease the use of cash-draining inputs and increase the use of those inputs which are not cash draining: $\partial k / \partial \theta > 0$ and $\partial x / \partial \theta < 0$. This will hold for some x and k where the inputs are complements, that is, where y_{xk} is positive but sufficiently close to zero. The derivative of total product with respect to θ represents the effect on supply of increasing the importance of immediate cash revenues. Increasing the deviation of optimal input decisions from strict profit maximization may either increase or decrease supply, again depending on the degree of substitutability of inputs:

$$\frac{\partial y}{\partial \theta} = y_k \frac{\partial k}{\partial \theta} + y_x \frac{\partial x}{\partial \theta}.$$

Note that for $\theta = 0$, one can solve the above comparative static results to find that the sufficient condition for output to increase is given by $y_{xk} < w/i y_{kk}$.

The marginal conditions in (8a), (8b), and (8c) can be viewed more generally by writing the first-order conditions, given by expressions (8), as

$$(9a) \quad \frac{\partial y}{\partial x} = \rho_1 \frac{w}{\mu}$$

$$(9b) \quad \frac{\partial y}{\partial k} = \rho_2 \frac{i}{\mu},$$

where the ρ 's reflect a general systematic rule of behavior, with strict profit-maximization as a special case ($\rho_i = 1$). From expressions (9), the ρ 's are related through their mutual dependence on θ , μ , and p_d : $\rho_1 = \rho_2(1 + \theta\mu/p_d)$. These ρ 's are termed adjustment factors. For practical purposes, we take the ρ 's as functions of certain observable variables, which are not simultaneously also choice variables.⁷ For example, in the empirical section that follows, it is through government programs' influence on these adjustment factors that one determines the effectiveness of programs at reducing the effect of risk of farm failure.

A Special Case: An Overproduction Trap

This section considers a specific model of a producer's behavior under risk of forced adjust-

⁷ This approach is similar to that taken by Lau and Yotopoulos in their investigation of relative production efficiency. Lau and Yotopoulos take the ρ 's as constants over time for a given type of "farm." Their data uses averages of individual farms as a given size and region.

ment. This example is presented in two parts. In the first, the producer avoids risk by expending greater effort in producing a cash income in order to increase the likelihood of covering the minimal necessary expenditures to retain farm ownership. The purpose here is to show the possibility of a backward-bending supply curve, the discrete jumps in supply that may occur at various level of expected price, and the conditions under which a reduction in price variance reduces supply and improves productive efficiency.

The second part of this example considers market equilibrium. A stable, long-run equilibrium is defined where the probability of forced adjustment is zero and where expectations are rational. This example shows that rational-expectations equilibria can arise where the producer is permanently taking the risk of failure into account when making production decisions (i.e., producing "inefficiently"), but the observed probability of failure is zero.

We will term the producer behavior where output expands in response to falling output price survival mode and the equilibrium where supply is backward bending with a zero probability of failure, a survival mode equilibrium. If an outside observer finds a small probability of forced adjustment or failure, this does not necessarily imply the farmer is out of survival mode and making decisions based on maximizing farm-income-derived utility. Causality may flow the other way. This example demonstrates a case in which zero probability of failure exists, the farmer is rationally producing "inefficiently" (i.e., contrary to strict profit maximization), and supply increases with a fall in expected price. The important point is that the probability of failure is low (in this case, zero) because the farmer produces in this (strictly defined) inefficient manner.

Furthermore, if a survival-mode equilibrium exists with an inelastic demand curve, then another rational-expectations equilibrium with productive efficiency also exists. This two-equilibria condition illustrates an overproduction trap. That is, as a result of overproduction, market conditions are such that overproduction is optimal for individual farmers. Moreover, if in concert farmers reduced production to that of a conventional equilibrium, no individual would have incentive to expand. With an elastic demand curve, a survival-mode equilibrium may exist, but this cannot strictly be called a trap because only a single rational-expectations equilibrium exists.

Suppose there is one competitive farmer in the market producing some level of commodity, y , out of effort, e , and receiving some level of price, p . Production is random, and, therefore, so is price. Although price depends on quantity, the farmer acts as if he had no influence over its probability distribution function. Let the utility function from farming be the sum of the goods consumed out of farm revenues, py , over some minimal expenditure f , and leisure time: $U = (py - f) + (1 - e)$. This utility function implies a constant marginal utility of income and exhibits Arrow-Pratt risk neutrality. Random production is given by $y = 2e^{1/2}\varepsilon$, where ε is a random term taking on two values (associated say with bad and good weather), $\varepsilon = \varepsilon_l$ or $\varepsilon = \varepsilon_h$ ($\varepsilon_l < \varepsilon_h$), with equal probability. Let \bar{y} represent the expected or average yield, that is, $\bar{y} = 2e^{1/2}$. The demand curve is of constant elasticity $p = ay^{-b}$. Therefore, random revenues can be written as

$$(10) \quad py = a(2e^{1/2})^{-b} 2e^{1/2}\varepsilon^{1-b}$$

This problem with output and price random can be reduced to a simpler conceptual problem with only one source of randomness. It is conceptually easier to redefine price as having an expected value of $\mu = a\bar{y}^{-b}$ with multiplicative error of $u \varepsilon^{1-b}$, where without loss of generality $E[u] = 1$. Now the new random term u can take on two values associated with the two values of ε : $u = u_l$ or $u = u_h$ ($u_l < u_h$). The competitive farmer acts as if he cannot influence expected price; the farmer views revenues as $py = \mu u 2e^{1/2}$. The failure condition is where $py < f$, or where $e < (f/2\mu u)^2$.

To summarize, the farmer's objective function is given by

$$(11) \quad V = \frac{\bar{U}(e) - I}{1 - \beta(1 - \pi)},$$

where $\bar{U}(e)$ is the expected farm-derived utility, I the alternative utility of leaving farming, β the personal discount rate, and π the probability of failure. The probability of failure may take on three values depending on the chosen level of output. If the farmer expends so little effort that even at the highest possible price cash receipts do not cover minimal expenses, then the probability of failure is one. If, on the other hand, the farmer expends enough effort so that even at the lowest possible price receipts cover minimal expenses, then the probability of failure is zero. For levels of effort in a middle range, either the farmer fails or he does not, depending on

the outcome of price. The probability of failure is the probability of the low price being realized, which in this example is 50%.

An important aspect of this model to note is that, except at the boundaries of these regions, expending incrementally any more or less effort will not affect the farmers probability of failure; it will only increase or decrease his benefits from farming. One may set, without loss of generality, the nonfarm utility level to zero, $I = 0$.

Consistent with the condition cited above, the objective function may then be represented in one of three ways, depending on whether the farmer's effort makes the probability of failure one, one-half, or zero:

$$(12a) \quad \mu 2e^{1/2} + (1 - e) - f \quad \text{if } \pi = 1, \text{ or } e^{1/2} < \frac{f}{2u_h\mu}$$

$$(12b) \quad V = \frac{\mu 2e^{1/2} + (1 - e) - f}{1 - .5\beta} \quad \text{if } \pi = .5, \text{ or } \frac{f}{2u_h\mu} \leq e^{1/2} < \frac{f}{2u_l\mu}$$

$$(12c) \quad \frac{\mu 2e^{1/2} + (1 - e) - f}{1 - \beta} \quad \text{if } \pi = 0, \text{ or } \frac{f}{2u_l\mu} \leq e^{1/2}$$

Figure 1 shows one possible set of values of the function V . The concave lines show the values of the objective function for given levels of π . That is, the lowest curve represents expression (12a) over all values of effort, the middle curve represents expression (12b), and the highest curve represents expression (12c). The heavily drawn lines show the objective function taking π into account for a specific pair of price outcomes; these lines represent the set of possible choices of effort open to the farmer. For the case shown in the figure, optimal effort is where

$$(13) \quad e^* = \left[\frac{f}{2u_l\mu} \right]^2;$$

that is, optimal effort is chosen such that the probability of failure is zero. This illustrates the danger of taking observed probabilities of failure as exogenous to the farmer's supply decisions. Here there is no chance of failure, but the level of effort chosen is away from the point of "efficiency," ($e = \mu^2$), which would be optimal if risk of failure were irrelevant to decisions.

Optimal effort is conditional on regions in which expected price may fall. Optimal supply over expected price is graphically illustrated in figure 2. The farmer may not choose to be on a portion of the supply curve where $\pi = 1/2$, instead either choosing a supply where $\pi = 0$,

or where $\pi = 1$. The downward sloping portions of the supply curve are the regions of expected price where the farmer is in survival mode. The effect of eliminating price variability, abstracting from equilibrium effects, is illustrated in figure 2. Consider an expected price of $\bar{\mu}$. Eliminating variance (i.e., setting $u_l = u_h$) yields an optimal supply of $\bar{y}^* = 2\bar{\mu}^2$.

Now consider the market equilibrium. A stable, long-run equilibrium is an expected market price, μ_e , a level of effort e_e^* , and an expected supply \bar{y}_e^* , such that the number of producers is constant (i.e., $\pi = 0$), and where

$$(14a) \quad \mu_e = a[\bar{y}_e^*]^{-b}$$

$$(14b) \quad \bar{y}_e^* = 2e_e^{*1/2}$$

$$(14c) \quad e_e^* = e^*(\mu_e).$$

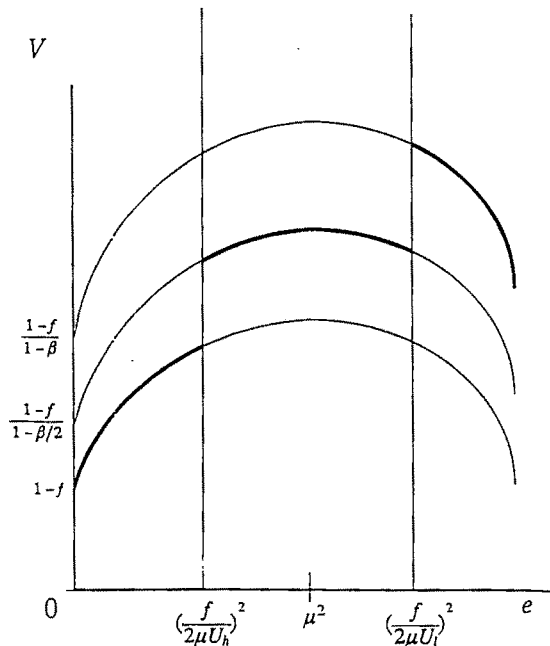


Figure 1. Values of the objective function over effort

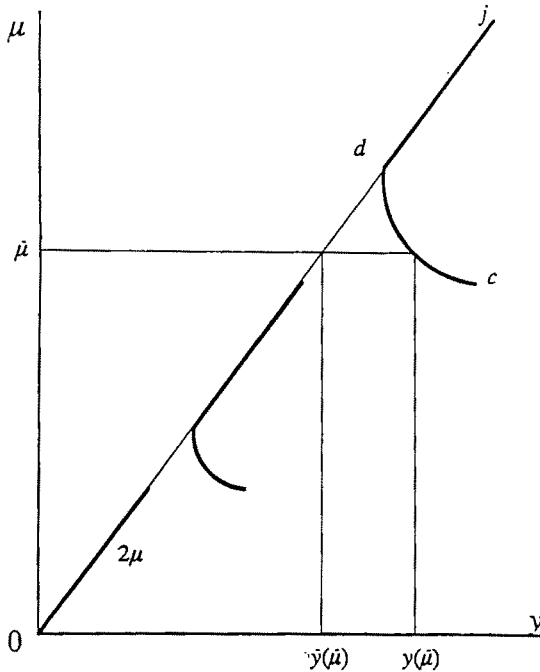


Figure 2. Optimal supply over expected price

Equation (14a) represents market equilibrium, (14b) represents the optimal expected supply given the optimal choice of effort, and (14c) represents the rational-expectations equilibrium where optimal effort is consistent with the equilibrium expected price. Stability implies that equilibrium price falls along the segments cd and dj on the supply curve in figure 2. The most interesting case is where the farmer is in survival mode but $\pi = 0$; that is, on segment cd of figure 2.

For a survival mode equilibrium to exist, expected price must be such that

$$(15) \quad \mu_e < \left[\frac{f}{2u_l} \right]^{1/2}.$$

In this case

$$(16) \quad \mu_e = a^{1/1-b} \left[\frac{f}{u_l} \right]^{-b/1-b}.$$

For an inelastic demand curve ($b > 1$), the conditions for survival-mode equilibrium are given by

$$(17) \quad \frac{f}{u_l} < 2^{1-b/1+b} a^{2/1+b}.$$

In fact, for all cases of demand, the above

condition is simply that which provides for the standard (nonsurvival-mode) equilibrium. That is, if a survival-mode equilibrium exists, then a conventional one does also. A conventional equilibrium exists where

$$(18) \quad \mu_e = (2a)^{1/1-b}.$$

An elastic demand may also yield a stable, survival-mode equilibrium; but, if one does not exist, then a conventional one would not exist as well. [The inequality in expression (18) above is reversed.] This leads to the idea of an overproduction trap. For both inelastic and elastic demands, survival-mode equilibria may exist. But only in the former case is one justified in using the term trap. In the latter case, survival mode arises only because of the objectives of farmers. In the inelastic-demand case, survival mode exists because of the objectives of farmers and the accident of expectations consistent with survival mode. Farmers are trapped by their rational expectations; without altering farmers' objectives, a conventional equilibrium may be attained.

Empirical Application

The preceding sections have discussed the influence of human-capital asset fixity on production decisions. This section empirically addresses farm production based on the conceptual model. First, even under risk neutrality the usual assumptions underlying the use of cost and profit functions are inapplicable in this case because the marginal-product-equals-price rule does not hold. Therefore, even if one used only disaggregated data, the conventional correlations of cost shares, for example, with factor prices would not represent production technology as standard application of duality theory would suggest. The difficulty is that marginal products are set equal to effective prices, which are unobserved. These effective prices are the observed prices adjusted by other factors reflecting the influence of the discrete difference between on-farm and off-farm utility and the probability of failure.

In standard applications, there are estimable equations for a production function (or cost, or profit) function and each marginal condition. If there are n choice variables with associated prices, the standard application would have $n + 1$ equations from which to estimate the parameters representing the production technology (n input demands, one output supply). One may solve n input choice variables for n prices, all of which

can be observed. In the nonstandard case described above, however, one can solve only for the choice variables in terms of the n prices and two price-adjustment factors, $1/(1 + \theta)$ and $(1 + \theta \mu/p_d)/(1 + \theta)$ from equations (8a) and (8b). Nevertheless, the marginal rates of technical substitution between inputs with common adjustment factors are dependent only on observable prices. Therefore, at best, $n - 2$ choice variables can be solved in terms of observable prices and two inputs associated with different adjustment factors. And one can obtain $n - 1$ equations from which to estimate production function parameters ($n - 2$ input demands, one output supply). This is not surprising, because this model introduces at least two additional unknowns into the choice problem. Under risk neutrality, besides the parameters defining technology, the true cost ($\bar{U} - I$) of leaving farming in any period is unobserved, as is the true probability of failure as a function of farm decisions. In order to estimate the production function parameters, unconditioned on the on-farm/off-farm utility difference and the failure probability, one must allow for the influence of these two additional unknowns through some similar number of observables.

To illustrate the conceptual and theoretical models above, we examine corn production in Illinois. This study estimates a per-acre production function, assuming a particular production function and utilizing the plausible restrictions on parameter estimates implied by the behavior model of a profit-maximizing farmer. Of course, other models also could result in production inefficiency, defined here as a wedge between marginal products and observed prices. The empirical analysis tests whether the adjustment factors driving a wedge between marginal products and prices, move in the direction implied by the conceptual model of input decisions under risk. Specifically, the empirical model tests whether or not input choices approach productive efficiency as farm size increase and over time (as farming is hypothesized to grow more integrated into the larger economy). Additionally, in recognition of the widespread farmer use of government programs and their influence on risk and farm decisions, the analysis examines to what degree programs affect the deviation of input choice from productive efficiency.

The empirical analysis takes four inputs (in per acre amounts) to a Cobb-Douglas production function. Consistent with the conceptual model, the inputs are separated into those that must be paid for immediately, fertilizer and hired

labor, and those that may go unpaid, farm family labor and physical capital, represented by machinery use. In order to reduce the scope of the problem, the possible influence of other crops is ignored—both in the usual joint-production sense, and, more importantly, in the sense of choosing a portfolio. The data are taken from *Summaries of Illinois Farm Business Records* from 1971 to 1979. This period was regarded as relatively prosperous for midwestern corn producers, with few downturns. Indeed, parts of this period were considered “boom times,” with many producers taking on greater debt consistent with a decreased perception of risk of farm failure. The period’s comparative prosperity, in marked contrast to the following years of the early 1980s, would tend to work against observing any influence of failure risk on input decisions. The data are averages of farm characteristics, production levels, factor expenses, and labor employed for farms in particular size ranges and regions. Price data are from the *Summaries*, the Department of Agriculture’s *Agricultural Statistics*, and the Commodity Research Bureau’s *Commodity Year Book*.

Yield is represented by Y , and input levels by X_i ; where $i = 1$ denotes fertilizer, $i = 2$ hired labor, $i = 3$ family labor, and $i = 4$ machinery. The W_i represent prices associated with the inputs, and the α_i represent associated elasticities to be estimated. We assume risk neutrality of farmers and that output price is independent of output. The March quote of the December futures contract for corn represents the expected output price. Denote this expected price by μ . The model also includes as shift variables the amount of corn acreage (Z_1) to account for a scale effect, and the average soil quality of the farm (Z_2). (See the *Summaries* for a definition of this variable.)

The production function is represented as

$$(19) \quad \ln Y = \tilde{\alpha}_0 + \sum_{i=1}^4 \alpha_i \ln X_i + \sum_{j=1}^2 \beta_j \ln Z_j.$$

From the general first-order conditions given by expressions (9a) and (9b), the marginal conditions for optimization are

$$(20a) \quad \frac{\mu Y}{W_i X_i} \alpha_i = \rho_1 \quad i = 1, 2,$$

$$(20b) \quad \frac{\mu Y}{W_j X_j} \alpha_j = \rho_2 \quad j = 3, 4,$$

where ρ_k , ($k = 1, 2$) represents the unknown adjustment factors associated with the two types of inputs. The marginal product restrictions within each input group are

$$(21a) \quad X_2 = \frac{\alpha_2}{\alpha_1} \frac{W_1}{W_2} X_1,$$

$$(21b) \quad X_4 = \frac{\alpha_4}{\alpha_3} \frac{W_3}{W_4} X_3.$$

Therefore, one can specify the system of equations

$$(22a) \quad \ln Y = \alpha_0 + \alpha_1 \ln X_1 + \alpha_2 \ln \frac{W_1 W_1}{W_2} + \alpha_3 \ln X_3 \\ + \alpha_4 \ln \frac{W_3 X_3}{W_4} + \beta_1 \ln Z_1 + \beta_2 \ln Z_2 + u_1$$

$$(22b) \quad \frac{W_2 X_2}{\mu Y} = r_1 \frac{W_1 X_1}{\mu Y} + u_2$$

$$(22c) \quad \frac{W_4 X_4}{\mu Y} = r_3 \frac{W_3 X_3}{\mu Y} + u_3,$$

where $r_1 = \alpha_2/\alpha_1$, and $r_3 = \alpha_4/\alpha_3$. These three equations may be estimated using a seemingly unrelated regression technique that imposes nonlinear restrictions on the coefficients.⁸

The estimates of the production function elasticities are reported in table 2. The estimated production elasticities with respect to the inputs are positive and of plausible magnitude. One noteworthy result is the large difference between the elasticity of family labor and that of hired labor. Another result to note is the sign and magnitude of the coefficient on acreage under corn, indicating a positive scale effect. The per-acre production function, however, exhibits decreasing returns.

Of particular interest is how the adjustment factors (ρ_i) are influenced by non-input variables. In order to investigate this, assume that the adjustment factors may be represented in the following manner:

$$\rho_i = \frac{\mu Y}{W_i X_i} \alpha_i = a_i + b_i S + c_i T + d_i D \\ + e_i M + u_{1i} \quad i = 1, 2$$

Table 2. Estimates of Production Elasticities

Variable	Coefficient	Estimate	<i>t</i> -statistic
Constant	α_0	3.0282	6.0490
Fertilizer	α_1	0.1015	1.4472
Hired labor	α_2	0.0174	1.4367
Family labor	α_3	0.1407	4.3126
Machinery	α_4	0.2280	4.3200
Corn acreage	β_1	0.2555	3.5311
Soil quality index	β_2	0.0085	11.773

$$\rho_2 = \frac{\mu Y}{W_j X_j} \alpha_j = a_2 + b_2 S + c_2 T + d_2 D \\ + e_2 M + u_{2j} \quad j = 3, 4,$$

where S represents farm size in gross acreage, T a time index, D the percent tillable land under program diversion, and M the proportion of off-farm income to total income.

From the discussions in the previous sections, one expects that as farm size grows the adjustment factors approach one, falling ($b_1 < 0$) in the case of fertilizer and hired labor ($i = 1, 2$), and increasing ($b_2 > 0$) in the case of family labor and machinery ($j = 3, 4$). These effects reflect that the fixed costs of farming which must be covered in order for the farmer to remain in agriculture do not grow in proportion with farm size. The larger operations relative to the smaller ones spread these fixed costs over a greater number of acres, and this optimally sets marginal products closer to observed factor prices. The time index, T , reflects an overall improvement in alternatives for farm family labor; growing integration of the farm sector and the larger economy, and other structural changes that promote stricter profit maximizing (that is, $c_1 < 0$ and $c_2 > 0$).

One variable of special concern indicating the integration of farm labor into other sectors is the relative importance of off-farm income to the total farm family income. As the family grows less dependent on uncertain production-based returns and more dependent on sure, off-farm sources, the farmer places less weight on avoiding poor crop revenues ($e_1 < 0$ and $e_2 > 0$). This effect would best be measured by the proportion of off-farm income relative to total income for specific farm families. This set of data was unavailable. Instead the model takes the proportion of total nonfarm income earned by U.S. farmers relative to total income from both farm and nonfarm sources. Thus, the data reflect only

⁸ One potential difficulty with the SUR approach is the possible simultaneity of the representative inputs and the level of production.

aggregate changes in off-farm income to commercial enterprises.⁹

The program variable, D , reflects farmer's response to government payments. Three effects from government programs can be identified. First, programs may reduce the risk of failure and thereby promote efficiency. Second, programs may increase the opportunity cost of leaving farming, at the same time reducing failure risk, and thereby promote inefficiency. Third, programs may increase the effective output price through target prices. Despite these three distinct effects, one may be able to detect which has the greater influence by examining the pair of coefficients d_1 and d_2 . If the first effect predominates, then the signs of the coefficients on D are expected to be positive for family labor and machinery ($d_2 > 0$), and negative for fertilizer and hired labor ($d_1 < 0$). If the second effect predominates, then the signs on this coefficient will be reversed ($d_1 > 0$ and $d_2 < 0$). Finally, if the third effect predominates, then the signs on both coefficients will be negative for both equations ($d_1 < 0$ and $d_2 < 0$).

Using the first-order conditions, specify four additional equations:

$$(23a) \quad \frac{\rho_1}{\alpha_1} = \frac{\mu Y}{W_1 X_1} = \frac{a_1}{\alpha_1} + \frac{b_1}{c_1} S + \frac{c_1}{\alpha_1} T + \frac{d_1}{\alpha_1} D + \frac{e_1}{\alpha_1} M + \frac{u_{11}}{\alpha_1}$$

$$(23b) \quad \frac{\rho_1}{\alpha_2} = \frac{\mu Y}{W_2 X_2} = \frac{a_1}{\alpha_2} + \frac{b_1}{\alpha_2} S + \frac{c_1}{\alpha_2} T + \frac{d_1}{\alpha_2} D + \frac{e_1}{\alpha_2} M + \frac{u_{12}}{\alpha_2}$$

$$(23c) \quad \frac{\rho_2}{\alpha_3} = \frac{\mu Y}{W_3 X_3} = \frac{a_2}{\alpha_3} + \frac{b_2}{\alpha_3} S + \frac{c_2}{\alpha_3} T + \frac{d_2}{\alpha_3} D + \frac{e_1}{\alpha_3} M + \frac{u_{23}}{\alpha_3}$$

$$(23d) \quad \frac{\rho_2}{\alpha_4} = \frac{\mu Y}{W_4 X_4} = \frac{a_2}{\alpha_4} + \frac{b_2}{\alpha_4} S + \frac{c_2}{\alpha_4} T + \frac{d_2}{\alpha_4} D + \frac{e_1}{\alpha_4} M + \frac{u_{24}}{\alpha_4}$$

These equations may be estimated using SUR with linear cross-equation restrictions. The restrictions are obtained from estimates of the production elasticities. For example, the coefficients on the national off-farm income variable

for the cash-draining inputs, say e'_1 and e''_1 , are restricted such that $e'_1 = \alpha_2/\alpha_1 e''_1$.¹⁰

The estimates of the coefficients for the adjustment-factor equations are reported in table 3. The coefficients on farm size in the adjustment-factor equations support the conclusions of the conceptual model. Over farm size the adjustment factor falls for fertilizer and hired labor and increases for family labor and machinery. The coefficients are both significantly different from zero with a degree of confidence greater than 95%. The coefficients are also of the same magnitude, although of different sign (as expected), suggesting that the effect of farm size on the deviation from strict profit maximization is roughly symmetric with respect to decisions regarding both types of inputs. Similarly, the coefficients on the variable measuring the importance of off-farm income are of expected sign. The greater the proportion of off-farm income, the less the marginal products of fertilizer and hired labor exceed their observed factor prices and the less the marginal products of family labor and machinery fall short.

The coefficient on the off-farm income variable, however, is insignificantly different from

⁹ A reviewer notes that off-farm income in some cases could be approximated by aggregate data by sales classes. Such class data, however, are not immediately applicable here. The farm sizes here for the most part are of the commercial class, although not exclusively. Moreover, the farm sizes used change over time. The results should be interpreted in this light, specifically as if the relative off-farm income across farms was a constant, and the data reflect only common changes in levels.

¹⁰ There are two noteworthy potential problems with this estimation method. First the restrictions rely on estimates of the production elasticities. Interpretation of the standard errors must rely on asymptotics and caution is advised. Second, the error terms in these four equations above may not be independent of the errors in the equations representing the production relations. Ideally, a larger seven-equation system would be estimated, production relations and adjustment factors together.

Table 3. Coefficient Estimates for the Adjustment Factors

Variable	Coefficient	Estimate	t-statistic
Cash-draining inputs			
Constant	$a'_1 = \frac{a_1}{\alpha_1} = \frac{\alpha_2}{\alpha_1} a''_1$	17.639	5.580
Gross acreage	$b'_1 = \frac{b_1}{\alpha_1} = \frac{\alpha_2}{\alpha_1} b''_1$	-0.0095	-10.972
Time index	$c'_1 = \frac{c_1}{\alpha_1} = \frac{\alpha_2}{\alpha_1} c''_1$	0.5189	3.549
Percent tillable land	$d'_1 = \frac{d_1}{\alpha_1} = \frac{\alpha_2}{\alpha_1} d''_1$	0.2386	2.838
Ratio off-farm to total income	$e'_1 = \frac{e_1}{\alpha_1} = \frac{\alpha_2}{\alpha_1} e''_1$	-2.5203	-0.418
Non-cash-draining inputs			
Constant	$a'_2 = \frac{a_2}{\alpha_3} = \frac{\alpha_4}{\alpha_3} a''_2$	-6.457	-2.592
Gross acreage	$b'_2 = \frac{b_2}{\alpha_3} = \frac{\alpha_4}{\alpha_3} b''_2$	0.0133	19.491
Time index	$c'_2 = \frac{c_2}{\alpha_3} = \frac{\alpha_4}{\alpha_3} c''_2$	-0.2959	-2.566
Percent tillable land	$d'_2 = \frac{d_2}{\alpha_3} = \frac{\alpha_4}{\alpha_3} d''_2$	-0.3298	-4.975
Ratio off-farm to total income	$e'_2 = \frac{e_2}{\alpha_3} = \frac{\alpha_4}{\alpha_3} e''_2$	24.323	5.115

ing off-farm income affects the utilization of the different types of inputs in an asymmetric way. Off-farm income lessens the overuse (relative to strict profit maximization) of family labor and machinery but has little if any effect on discouraging the underuse of the other inputs.

The time index represents all variables other than off-farm income that would measure the degree of integration of farming with the rest of the economy. The coefficients on this composite variable suggest that, for a given farm size, dependence on farm income, and degree of program participation, production decisions have been deviating further from pure profit maximization. These results are somewhat surprising in light of accepted wisdom that agriculture is in transition to greater integration with the rest of the economy. First, recall that the data are for a fairly short period of time, between 1971 and 1979. Second, the results for the time variable are supported at least in part by Vasavada's work on the measurement of excess inputs. Using a dynamic adjustment model, he concludes that for aggregate levels of labor and capital,

surpluses have shown a marked tendency to decline over a longer period of time (since 1948). During the 1970s, however, his results demonstrate a decline followed by an upswing in input-surplus indices. Troughs occur in 1972 for capital and 1974-75 for labor.

Finally, the coefficients on the program variable are of opposite signs for the two adjustment-factor equations and support the conclusion that government programs exacerbate the deviation from production efficiency.

Concluding Comments

The analytical and empirical results of this paper offer some insight into farmer behavior under risk of failure. A farmer cannot purchase complete insurance against such a risk. In addition to other responses (e.g., crop insurance, access to credit reserves), the farmer would seek to mitigate against this risk by deviating in his production decisions from what is optimal from a simple expected-profit-maximizing case. Production factors with immediate cash outlay tend to have higher effective prices than without the risk because part of their cost must be measured in the contribution to increasing the probability of failure. The marginal products of these factors are set higher than observed prices would optimally warrant. Conversely, factors that may be delayed in cash expenditure tend to have lower effective prices for the opposite reason, and their marginal products are set lower than observed prices warrant.

Factors of the last type are of particular interest because their contribution to aggregate capacity may be of greatest significance. Farm-operator labor, or farm family labor, tends to be overutilized at the expense of non-cash-generating alternatives (e.g., leisure time). Improving opportunities for off-farm income would encourage the farmer to use operator labor in a similar manner to hired labor. Physical capital, including aspects of land quality and long-term productivity, owned by the farmer is treated in the same way. The conceptual analysis suggests that during periods when farmers face higher probability of failure and the difference between on-farm and off-farm utility is larger, farmers would tend rationally to increase the deterioration of their resources.

This paper also offers an empirical investigation of corn production in Illinois. This is done in order both to demonstrate the applicability of the conceptual model and to substantiate certain

conclusions that can be drawn regarding the degree of deviation from simple profit maximization. Conventional estimation of dual functions, such as those of cost and profit, is unwarranted in the presence of risky adjustment costs. Nevertheless, the theory does admit certain restrictions to an estimable system of supply and factor demands, from which one can use output- and input-price data in the estimation of production technology. In addition to prices as explanatory variables, levels of a certain number of representative inputs (in this study, two) must be used with a corresponding reduction in the number of estimable equations. The estimation results indicate that larger farms deviate less from production efficiency than do smaller farms, where production efficiency is defined relative to strict profit maximization; and that the lesser the reliance on farm income as opposed to off-farm income, the greater the production efficiency. Over the period of time studied, however, farmers have been moving further from setting marginal products equal to observed prices.

[Received September 1988; final revision received June 1990.]

References

- Akerlof, G. A. "A Theory of Social Custom, of Which Unemployment May Be One Consequence." *Quart. J. Econ.* 94(1980):749-75.
- . "The Market for Lemons: Quality, Uncertainty and the Market Mechanism." *Quart. J. Econ.* 84(1970):488-500.
- Commodity Research Bureau. *Commodity Year Book*. Jersey City NJ: Commodity Research Bureau, various editions, 1971-80.
- Galbraith J. K., and J. D. Black. "The Maintenance of Agricultural Production During Depression: The Explanations Reviewed." *J. Polit. Econ.* 46(1938):305-23.
- Illinois Cooperative Extension Service. *Summaries of Illinois Farm Business Records (Commercial Farms)*. Annual reports 1971-79, published by the University of Illinois College of Agriculture.
- Johnson, G. L., and C. L. Quance, eds. *The Overproduction Trap*. Baltimore MD: Johns Hopkins University Press, 1972.
- Johnson, M. A., and E. C. Pasour. "An Opportunity Cost View of Fixed Asset Theory and the Overproduction Trap." *Amer. J. Agr. Econ.* 63(1981):1-7.
- Just, R. E., and D. Zilberman. "In Defense of Fence to Fence." *Dep. Agr. and Resour. Econ. Work. Pap.*, University of Maryland, Aug. 1988.
- Lau, L. J., and P. A. Yotopoulos. "A Test for Relative Efficiency and an Application to Indian Agriculture." *Amer. Econ. Rev.* 56(1971):94-109.
- Pyle, D. H., and S. J. Turnovsky. "Safety-First and Expected Utility Maximization in Mean-Standard Deviation Portfolio Analysis." *Rev. Econ. and Statist.* 52(1970):75-81.
- Robison, L. J., P. J. Barry, and W. G. Burghardt. "Borrowing Behavior Under Financial Stress by the Proprietary Firm: A Theoretical Analysis." *West. J. Agr. Econ.* 12(1987):144-51.
- Thompson, O. E., D. Gwynn, and C. Sharp. "Characteristics of Women in Farming." *California Agr.* 41(1987):16-17.
- U.S. Department of Agriculture. *Agricultural Statistics*. Washington DC, various editions, 1971-82.
- Vasavada, U. "Is U.S. Agriculture Overcapitalized?" *Dep. Agr. Econ. Work. Pap.*, University of Georgia, September 1986.

Valuing the Multidimensional Impacts of Environmental Policy: Theory and Methods

John P. Hoehn

A theoretically consistent framework is developed for valuing the multidimensional impacts of environmental policy. Conventional benefit estimates are shown to be biased because of the presence of substitution and complementarity effects in valuing policy impacts. Procedures are developed for implementing a valid framework. Consistent with theory, empirical results indicate significant substitution effects in valuing environmental conditions across different geographic regions.

Key words: air quality, benefit cost analysis, contingent valuation, environmental policy, national parks, nonmarket valuation.

Environmental management requires tradeoffs across different environmental services and different user groups. Management policy is inherently multidimensional. A typical policy change is not simply to provide more or less of a single environmental characteristic but a decision as to what package or set of characteristics to provide.

The multidimensional nature of environmental policy has important implications for valuing policy change. As policy changes the package of environmental services, individual users make substitutions across a changing opportunity set. These substitutions affect the value that individuals place on both the overall policy package as well as its components.

Natural resource economists have been aware of the impact of substitution on resource valuations for some time. In the early 1970s, Burt and Brewer began the development of methods for valuing the attributes of a system of recreation sites. Knetsch discussed the problem of valuing existing recreational facilities that are displaced by a new recreational investment. Rosenthal showed the empirical biases that result when substitute prices are omitted from a travel cost model.

John P. Hoehn is an associate professor, Department of Agricultural Economics, Michigan State University.

Support from the U.S. Environmental Protection Agency and the Michigan Agricultural Experiment Station is gratefully acknowledged.

The author benefited from the referees' comments and discussions with Alan Randall, Richard Carson, Robert Mitchell, James Oehmke, and John Stoll.

The valuation concepts developed in the context of local recreation sites have not been extended to the general problem of valuing the benefits and costs of a multidimensional environmental policy. Instead, conventional benefit cost procedures tend to proceed on a piecemeal basis. Each element of a multiple impact policy is evaluated independently, as if it were the only element to be changed by policy. The conventional approach ignores any substitution or complementarity effects that may occur. A total valuation is obtained by summing across the independent component valuations.

It is clear that the conventional benefit cost design is conceptually and empirically flawed. Hoehn and Randall show that conventional procedures overstate the net benefits of a large, multidimensional policy change. Hoehn and Randall find that the conventional approach actually misidentifies detrimental policies as potential Pareto improvements. The results of Braeutigam and Noll; and Majid, Sinden, and Randall indicate a significant empirical divergence between the conventional and valid valuation approaches.

This paper examines a valid framework for valuing the multidimensional impacts of environmental policy. The first section of the paper outlines both the conventional and valid valuation designs.¹ The second section examines the prospect of substitution and complementarity between policy components. The third section

¹ Hof et al. consider similar issues for a firm's cost structure.

derives procedures for implementing the valid design with contingent valuation. The final section estimates valid bid functions and tests the empirical significance of substitution in valuing regional environmental quality conditions.

Conventional and Valid Designs

This section examines the valuation of environmental impacts brought about by a change in public policy. Policy impacts include a range of environmental effects. Impacts may include changes in physical resource flows, characteristics of the legal and social environment, or altruistic services such as those leading to existence value.²

Valuing a Single Impact Policy

The value of an environmental change stems from the tradeoffs a household is willing to make between (a) a change in an environmental service and (b) other desirable goods and services. Analytically, these tradeoffs are summarized in terms of a household's expenditure function. The expenditure function is a basic tool in defining the Hicksian measures of value.

Household preferences across market goods, $x = (x_1, \dots, x_L)$, and environmental services, $q = (q_1, q_2)$, are described by a utility function, $u = u(x, q_1, q_2)$, that is strictly increasing, continuous, and strictly quasi-concave. Given income, m , market prices, p , and access to environmental services, q , the level of well-being attained by a household is described by an indirect utility function, $v(p, q_1, q_2, m)$. At initial income, m^0 , and initial environmental quality, $q^0 = (q_1^0, q_2^0)$, initial utility is $u^0 = v(p, q_1^0, q_2^0, m^0)$.

The household's expenditure function is

$$(1) \quad e(q_1, q_2, u) = \min_x px \quad \text{s.t. } u(x, q_1, q_2) \geq u$$

where p is left implicit. Given the assumptions on utility, $e(q_1, q_2, u)$ is strictly decreasing and convex in q (Mäler). At initial environmental quality, q^0 , initial income maintains initial utility, $m^0 = e(q_1^0, q_2^0, u^0)$.

The Hicksian compensating measure of benefit, HC , is the amount of income, paid or re-

ceived, that would leave a household at a pre-policy level of well-being at the post-policy level of environmental quality. For a single impact policy that changes q_1 from q_1^0 to q_1^1 , HC is

$$(2) \quad HC(q_1^1, q_2^0; q^0) = m^0 - e(q_1^1, q_2^0, u^0).$$

where $m^0 = e(q_1^0, q_2^0, u^0)$. Equation (2) is the mathematical structure that underlies the Bradford bid curve (Bradford; Brookshire, Randall, and Stoll).

A Bradford bid curve analogous to equation (2) is shown in figure 1. The bid curve has a value of zero at the initial level of environmental quality. For a policy change from q_1^0 to q_1^1 that is beneficial to a household, the expenditure required to sustain the initial level of well-being declines and HC is positive. In this case, HC measures a household's willingness to pay (WTP). For a detrimental policy such as the change from q_1^0 to q_1^2 , the expenditure required to sustain initial utility is greater after the policy change and HC is negative. HC times minus one is the willingness to accept compensation (WTA).

The Evaluation of a Multidimensional Policy

A multidimensional policy results in the simultaneous change of two or more environmental services. This section compares the structure of a conventional evaluation of a multidimensional policy with a valid design.

The conventional approach values a multidimensional policy by independent valuation and summation (IVS). IVS values each of the im-

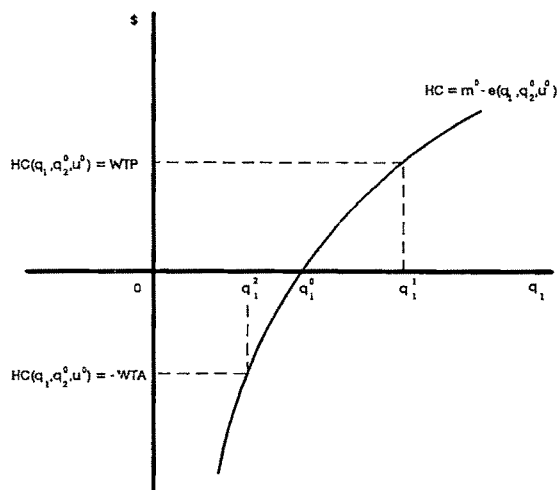


Figure 1. A Bradford bid curve

² Randall and Stoll discuss economic motivations for existence value.

pacts of a multidimensional policy as if each were a single impact policy. The independent valuations are added up to obtain an aggregate valuation.

For a two-element change from (q_1^0, q_2^0) to (q_1^1, q_2^1) , the *IVS* valuation is

$$(3.1) \quad IVS(q^1; q^0) = m^0 - e(q_1^1, q_2^0, u^0),$$

$$(3.2) \quad + m^0 - e(q_1^0, q_2^1, u^0).$$

The *IVS* approach fails to account for the aggregate impact of the policy change from q^0 to q^1 . Each of the policy impacts is evaluated with the other environmental parameter held constant at its initial level.

For a multiple impact policy, *HC* is the amount of income paid or received that would leave a household at the initial level of utility subsequent to the multiple impacts of policy. For the vector change from q^0 to q^1 , *HC* is

$$(4) \quad HC(q^1; q^0) = m^0 - e(q_1^1, q_2^1, u^0);$$

HC is unique and encompasses the impact of a multidimensional policy in a single, one-step valuation. In principle, contingent valuation can be used to implement equation (4) as a one-step, holistic valuation.

A disaggregated valuation is derived from equation (4) using a valuation path that begins at (q_1^0, q_2^0) and ends at (q_1^1, q_2^1) . The path may be either sequential or simultaneous. A sequential path shifts policy components one at a time from their initial to post-policy level. Using a sequential path that changes (q_1^0, q_2^0) to (q_1^1, q_2^0) , first, and shifts (q_1^1, q_2^0) to (q_1^1, q_2^1) , second, *HC* is

$$(5.1) \quad HC(q^1; q^0) = m^0 - e(q_1^1, q_2^0, u^0)$$

$$(5.2) \quad + e(q_1^1, q_2^0, u^0) - e(q_1^1, q_2^1, u^0).$$

The sequential valuation indicates how a valid disaggregate design accounts for the impact of the overall policy as well as its respective components. Line (5.1) values the change in q_1 with q_2 held constant at its initial level. Line (5.2) values the change in q_2 with q_1 at its post-policy level. Since the second term on the right-hand side of line (5.1) cancels with the first term in line (5.2), the sequenced valuation reduces to a quantity that is identical to equation (4)—the aggregate valuation.

Although the aggregate valuation is unique, the sequenced component valuations vary with the selected path. With a different valuation path—one that changes q_2 first and q_1 second—the component valuations in equation (5) would

be conditioned on a different level of the other environmental service. A change in conditioning variables changes the sequenced valuations.³

A simultaneous path values each policy component as the overall policy vector changes from q^0 to q^1 . A disaggregate valuation corresponding to a simultaneous path is⁴

$$(6.1) \quad HC(q^1; q^0) = \int_{q^0}^{q^1} [\partial e(q_1, q_2, u^0) / \partial q_1] dq_1$$

$$(6.2) \quad + \int_{q^0}^{q^1} [\partial e(q_1, q_2, u^0) / \partial q_2] dq_2,$$

where lines (6.1) and (6.2) each evaluate a derivative of the expenditure function, $\partial e(q_1, q_2, u^0) / \partial q_i$, $i \in \{1, 2\}$, as the overall policy shifts from its initial to its post-policy level. With the simultaneous path, both the aggregate and the component valuations are unique.

The valid component valuations may be estimated using either contingent valuation or a demand-based, weak complementarity method such as the travel cost or hedonic approach. A sequenced contingent valuation presents a sequence of changes and elicits a valuation at each step of the sequence. A simultaneous contingent valuation elicits the aggregate valuation, first, and then asks a respondent to partition the aggregate valuation into portions that are relevant to the respective components. Travel cost or hedonic demands are substituted directly into equations (5) or (6) by using the mathematical relationship between the expenditure function and weakly complementary demands.⁵

For a small number of policy impacts, the relationship between *HC* and *IVS* depends upon the extent of substitution, independence, or complementarity between policy components. Elements of q are substitutes, independent, or complements in valuation as the marginal valuation of q_i decreases, remains constant, or increases with an increase in q_j ; that is, as $\partial^2 HC / \partial q_i \partial q_j = - \partial^2 e / \partial q_i \partial q_j$ is negative, zero, or positive.

IVS overstates the valid valuation if environmental services are substitutes. With substitutes, line (5.1) is equal to (3.1) but line (5.2)

³ Conditioning variables may be viewed as part of the valuation context. The valuation context or frame plays a key role in psychological theories of valuation (Fischhoff and Furby).

⁴ The definition of a line integral leads directly to the simultaneous path.

⁵ Weak complementarity is detailed in Bockstael and Kling.

is less than line (3.2). Substitution implies that both lines in equation (6) are less than their counterparts in equation (3). With complementarity, the opposite effect occurs and *IVS* understates the aggregate and component valuations. *IVS* is equal to *HC* only if q_1 and q_2 are independent.⁶

WTP and *WTA* are each affected differently by substitution and complementarity. If a policy change is beneficial ($q_i^1 > q_i^0$), *HC* equals *WTP* and the impact of substitution and complementarity on *WTP* is identical to the impact on *HC*. If a policy change is detrimental ($q_i^1 < q_i^0$), *WTA* equals $-HC$. Because of the change in sign, the impact of substitution and complementarity on *WTA* is just the opposite of their effect on *HC*. With substitution, the *IVS* measure of *WTA* is less than the valid measure of *WTA*.

These effects on *WTP* and *WTA* result in distorted benefit cost outcomes when the *IVS* design is used. Suppose that, consistent with a potential Pareto improvement, benefits are measured using *WTP*, costs are measured using *WTA*, and net benefit is aggregate *WTP* minus aggregate *WTA*. If environmental services are substitutes, *IVS* overstates *WTP*, understates *WTA*, and overstates the net benefit of policy change. If they are complements, *IVS* understates net benefit. The *IVS* and *HC* net benefit measures are equivalent only when environmental services are independent in valuation.

The Prospect of Independence in Valuation

A valid benefit cost design imposes significant costs on practical policy evaluation. In addition to the challenge of valuing single components, a valid benefit cost design accounts for interactions between different components. These additional costs could be substantially reduced if independence were either predictable or routine.

Two valuation contexts suggest a strong pros-

pect of at least approximate independence. The first is a large number case that characterizes national planning problems where numerous agencies independently evaluate a large number of different policy proposals. In this case, the effects of substitution and complementarity may cancel out as the number of policy components—and *IVS* errors—becomes large.

Hoehn and Randall show that *IVS* errors do not cancel out in the large number case. Resource scarcity within an economy imposes a bound on the valid measure of benefits while the *IVS* outcome is unbounded. *IVS* is certain to overstate net benefits as the number of evaluations grows large.

The second policy setting involves a small number of policy components. This case is relevant to the internal planning problem of a single agency. Although a single agency may not have control of the overall policy agenda, it may still seek an accurate evaluation of its own program. This section examines the small number problem.

Hypotheses regarding the small number case are derived from the structure of individual choice and valuation. Substitution and complementarity appear to result from specific linkages in producing a subsistence or leisure activity that uses environmental services as inputs. The absence of a specific linkage in consumption may lead to independence in valuation.

The absence of a direct linkage in consumption is routinely expressed by an additively separable utility function,

$$u = \sum_{h=1}^H u_h(x, q_h),$$

where $q_h = (q_{h1}, \dots, q_{hI})$, $h \in \{1, \dots, H\}$, represents $H \times I$ different environmental services (Green, Deaton and Muellbauer). Environmental services q_h are additively separable from services q_g , $g \neq h$. As differentiation shows, the marginal utility of component q_{hi} is unaffected by a change in q_{gj} since $\partial^2 u / \partial q_{hi} \partial q_{gj}$ equals zero, $i, j \in \{1, \dots, I\}$. Environmental services that enter the same utility component, u_h , are not additively separable because $\partial^2 u / \partial q_{hi} \partial q_{hj}$ equals $\partial^2 u_h / \partial q_{hi} \partial q_{hj}$, which may be positive, negative, or zero.

Contrary to simple intuition, theorem 1 shows that independence in consumption leads to strict substitution in valuation.

THEOREM 1. *Let the preferences of a household be represented by a utility function such*

⁶ The component valuation of the change from q_2^0 to q_2^1 is

$$e(q_1^1, q_2^0, u^0) - e(q_1^1, q_2^1, u^0) = \int_{q_2^0}^{q_2^1} [-\partial e(q_1^1, q_2, u^0) / \partial q_2] dq_2,$$

where r equals zero to obtain line (3.2) and r equals unity to obtain line (5.2). If q_1 and q_2 are substitutes, the integrand evaluated at q_1^1 is smaller than the integrand evaluated at q_1^0 . Thus, line (5.2) is smaller than line (3.2). If q_1 and q_2 are complements, then the opposite conclusion holds. If the two are independent, lines (3.2) and (5.2) are equal. A similar analysis applies to equations (3) and (6). Psychologists have discussed these effects as an embedding phenomenon (Kahneman and Knetsch).

that (a) q_{hi} is additively separable from q_{gj} and (b) q_{hj} and q_{hk} are not additively separable. Assume that the utility components, u_h , are concave in x and that environmental services and market goods are gross complements, i.e., $\partial^2 u_h / \partial q_{hi} \partial x_n > 0$. The following properties hold for environmental services with nonzero valuations: (a) environmental services that are additively separable in utility are substitutes in valuation and (b) environmental services that are not additively separable in utility may be substitutes, complements, or independent in valuation.⁷

The theorem demonstrates that substitution effects in valuation are the rule rather than the exception. Substitution occurs consistently across all valuation contexts and arises because of the mathematical structure of constrained optimization. Substitution effects in valuation are certain if there is no direct utility linkage between environmental services. Complementarity and independence are possible only if a complementary linkage outweighs the substitution effect induced by constrained choice.

Two routine valuation contexts lead to additive separability and strict substitution. The first involves spatially separated activities. A household's use of a ski site in region h may be unaffected by its use of a ski site in region g . Fishing opportunities in region h may be unaffected by boating opportunities in region g . In a household production framework, each regional activity may be viewed as a different production process, $u_h(\cdot)$, that is separable in consumption. Regional environmental conditions enjoyed in separable activities are strict substitutes in valuation.

The second source of dominating additivity stems from uncertainty and the expected utility property. For example, suppose $h \in \{1, 2\}$ and that the enjoyment of q_1 or q_2 depends on some probabilistic event. The probability of enjoying q_1 and being without q_2 is π_1 , the probability of enjoying q_2 but going without q_1 is $(1 - \pi_1)$. The expected utility property implies that

$$(7) \quad u = \pi_1 u_1(x_1, q_1) + (1 - \pi_1) u_2(x_2, q_2),$$

where x_h represents market goods consumed in state h . The expected utility property introduces additive separability across mutually exclusive outcomes. Environmental conditions associated with different uncertain outcomes are substitutes in valuation.

Environmental valuations often involve the concept of option price. Because they are derived within an analogous expected utility framework, option prices may be sensitive to substitution effects. IVS estimates of option prices may be sensitive to substitution effects. IVS estimates of option price are likely to overstate the valid measure of benefits.

Estimating the Benefits of Complex Policy

Accurate evaluation of policy benefits requires procedures for implementing a valid benefit cost design. As indicated above, a valid design may use either contingent valuation or demand-based techniques such as the travel cost and hedonic approaches. Demand-based approaches that encompass substitution and complementarity effects are relatively well understood (Deaton and Muellbauer). This section focuses on empirical procedures for use with contingent valuation.

Contingent Valuation Alternatives

Contingent valuation (CV) is in theory adaptable to any level of a valid valuation design. The issue with contingent valuation is largely one of experimental design—of how to develop adequate value information at the least cost. In some cases, the valuation problem is simple. For instance, if the composition of a policy is predetermined, a single valuation question may be formulated consistent with equation (4). There is no need to disaggregate and value each of the components.

The valuation problem is typically more complex. An agency may have identified a candidate set of policy instruments but be uncertain as to the composition of a final policy. Benefit information is used to evaluate both the aggregate policy as well as the components. In this case, one might attempt to identify all possible combinations of the candidate impacts and then evaluate each holistically according to equation (4). An all-inclusive approach is likely to be costly.

An econometric approach reduces the cost of exploratory valuation by extracting a maximum amount of information from a sample of value data. The first step is to draw a representative sample of policy scenarios and to elicit contingent valuations for each scenario as suggested by equation (4). The functional relationship underlying equation (4) is then estimated in a man-

⁷ A proof of theorem 1 is available upon request to the author.

ner analogous to those used with multiproduct cost functions (Baumol, Panzar, and Willig). The estimated relationship is a multidimensional Bradford bid function that gives the benefits associated with any combination and level of policy components.

There are three considerations in selecting a functional form for equation (4). First, the predetermined variables should enter the bid functions in a manner consistent with theory. This requires that the bid function be equal to zero when environmental quality is held constant at the initial level. Socioeconomic variables should also enter the functions so that the initial conditions hold. Variables representing environmental quality should be permitted to take on any sign.

Second, the convexity of the expenditure function in q implies a bid function that is concave in q . A concave bid function has a Jacobian that is negative semidefinite. Since the concavity restriction results from a convenient but not necessary utility assumption, a functional form should be flexible enough to allow either concavity or convexity. Such flexibility permits an empirical test of concavity.

Third, the function should admit either substitution, complementarity, or independence between different pairs of amenity services. Neither the Cobb-Douglas nor the constant elasticity of substitution (CES) function satisfies this second criterion. The Cobb-Douglas and CES forms both impose qualitatively uniform effects across different pairs of amenities.

Functional Forms

Admissible empirical forms for the bid function are obtained using the approach used to derive flexible production and cost functions. This approach approximates the true functional form with a second-order Taylor-series expansion (TSE) (Gallant).⁸ The TSE parameters that measure pairwise interactions may take on any sign and socioeconomic variables can be entered in a manner consistent with the initial conditions. Environmental services may be amenities or disamenities.

A quadratic bid function is obtained by applying the TSE directly to equation (4). Before

expanding, however, two alterations are appropriate. First, to remove an unobservable utility term, $u^0 = v(q^0, m^0)$ is substituted into (4). Second, consistent with the household production literature, the utility and expenditure functions depend on the socioeconomic characteristics of a household, \bar{s} . With these changes, the quadratic specification is

$$(8) \quad HC(q^1; q^0) = m^0 - c(q_1^1, q_2^1, q^0, s),$$

where $c(q_1^1, q_2^1, q^0, s) = e(q_1^1, q_2^1, \bar{s}, v(q^0, \bar{s}, m^0))$ and $s = (\bar{s}, m^0)$.

Define the n th policy scenario, $n \in \{1, \dots, N\}$, as a set of ordered pairs, $\{(q_1^w, q_1^z), (q_2^w, q_2^z)\}$, where w and z are from a real interval that includes zero and one.⁹ For a sample of R respondents and N policy scenarios, a TSE taken about the initial values q^0 and the mean value of s , \bar{s} , results in

$$(9) \quad hc_{rn} = \alpha_0 + ds_r' \alpha_s + dq_n \beta + ds_r' A dq_n + dq_n' B dq_n + \epsilon_{rn},$$

where hc_{rn} is the r th respondent's valuation of the n th policy scenario; α_0 is a constant; $ds_r = (s_r - \bar{s})$ is a vector of differences between the r th respondent's characteristics, s_r , and \bar{s} ; α_s is conformable coefficient vector; $dq_r = [(q_1^w - q_1^0), (q_2^z - q_2^0)]$ is a vector of differences between the r th post-policy level of environmental quality and the initial environmental quality level; $\beta = (\beta_1, \beta_2)$ is a two-element coefficient vector; A is a coefficient matrix conformable to ds_r and dq_n ; $B = [b_{ij}]$, $i, j \in \{1, 2\}$, is a matrix of environmental interaction coefficients; and ϵ_{rn} is a statistical error term. With the independent variables in their differenced form, the coefficient β_h represents the marginal value of a change in q_h from q^0 .

The error ϵ_{rn} may be correlated across observations if multiple responses are obtained from individual respondents. In this case, ordinary least squares are not minimum variance estimators. Efficient estimates are obtained using generalized least squares procedures developed for panel data.

The quadratic function (9) is adaptable to requirements of economic theory. First, the expected value of hc_{rn} is equal to zero at the initial quality level if α_0 and α_s are equal zero. These

⁸ TSE forms are suitable to a local approximation of the true functional form. Global properties of flexible functional forms may be undesirable (Caves and Christensen).

⁹ The variables w and z allow one to describe sampling designs that include more than one policy scenario. In previous sections, w and z are unity.

restrictions on α_0 and α_s can be tested using routine statistical procedures.

Second, the quadratic permits either a concave or convex bid function. The quadratic is concave (convex) if B is negative (positive) semidefinite. Concavity (convexity) requires that the diagonal elements of B are nonpositive (non-negative) and the principal minors of $|B|$ alternate in sign (are non-negative). For two quality elements, this means that b_{ii} is nonpositive for concavity, b_{ii} is non-negative for convexity, and $|B| = b_{11}b_{22} - b_{12}^2$ is non-negative in either case.

Third, the quadratic equation (9) admits substitution, independence, or complementarity. Differentiation shows that environmental services are substitutes, independent, or complements as the asymmetric, off-diagonal terms in B , b_{ij} , $i \neq j$, are negative, zero, or positive.

The structure of a bid function precludes the derivation of a transcendental logarithmic (translog) function directly from equation (8). With a translog, the logarithm of a zero valuation is undefined.¹⁰ However, a semilogarithmic quadratic (semilog) form can be derived by taking a TSE about the logarithms of the right-hand-side variables in equation (8). The semilog quadratic specification is

$$(10) \quad hc_{rn} = \rho_0 + dls'_r \rho_s + dlq'_n \gamma + dls'_r C dlq_n + dlq'_n D dlq_n + v_{rn},$$

where ρ_0 is a constant; dls_r is a vector of differences between the logarithms of the elements of s_r and the logarithms of the elements of \bar{s} ; ρ_s is a conformable coefficient vector; $dlq_n = [\log(q_1^n) - \log(q_1^0), \log(q_2^n) - \log(q_2^0)]$ is the vector of differences between the logarithms of the n th post-policy level of environmental quality and of the initial environmental quality level; $\gamma = (\gamma_1, \gamma_2)$ is a two-element coefficient vector; C is a coefficient matrix conformable to dls_r and dlq_n ; $D = [d_{ij}]$, $i, j \in \{1, 2\}$, is a matrix of environmental interaction coefficients; and v_{rn} is an error term. Statistical properties of v_{nr} are analogous to those of ϵ_{rn} .

Equation (10) is adaptable to the restrictions of theory. Here, $E(hc_{rn})$ is equal to zero at the initial quality level if the estimated ρ_0 is equal to zero. The convexity restrictions on D and the form of interaction between environmental quality levels are analogous, at the point of approximation q^0 and \bar{s} , to those of equation (9).

Empirical Tests for Interaction

In this section, both the quadratic and the semilog bid functions are estimated. The estimated parameters are used to test the statistical significance of interaction effects in valuation.

Data for the empirical analysis came from a contingent valuation experiment conducted in Chicago, Illinois, during 1980 and 1981.¹¹ The experiment asked two sets of respondents to value changes in visual air quality in two different regions of the United States. The first sample of respondents was drawn from residents of Chicago in the summer of 1980. These respondents were asked to value improvements in air quality for Grand Canyon National Park. Respondents were read a narrative and shown photographs that represented the baseline conditions and the feasible improvements in visual air quality. Independent valuations were obtained for Grand Canyon air quality improvements of 40%, 83%, and 376%.

The second sample of respondents was drawn from Chicago residents in the spring of 1981. Three of the valuation questions posed to this set of respondents are of interest. First, respondents were read a narrative and shown photographs that represented two improvements of 100% and 233% in visual air quality for the Chicago metropolitan area. Independent valuations were obtained for each of these prospective improvements. Respondents were also read a narrative and shown the photographs representing the 83% improvement for the Grand Canyon. A final question valued a joint policy that proposed a 100% improvement for Chicago and an 83% improvement for the Grand Canyon.

By theorem 1, substitution effects are expected between the Chicago and Grand Canyon valuations. For a typical Chicago resident, future use of Grand Canyon improvements is likely to involve substantial uncertainty. Enjoyment of regional air quality conditions is also separated by both space and time. Both uncertainty and the spatial dimension suggest a hypothesis of substitution effects in valuation.

Before analyzing the data, valuations were screened to remove protest bids. Protest bids are zero valuations submitted by respondents who objected to the valuation procedure (Mitchell and Carson).

Table 1 lists the value data and sample characteristics for the 1980 and 1981 surveys. The independent valuations of Grand Canyon air

¹⁰ A translog form was derived from the dependent variable $(m^0 - HC)/m^0$. The translog proved intractable in testing and inferior in specification tests.

¹¹ Data collection procedures are detailed in Tolley and Fabian.

Table 1. Sample Value Data and Characteristics

Sample and Variable Names ^a	Mean	Standard Error
1980 Sample ^b		
Value of 40% improvement, G. Canyon (129)	\$58.8	1.0
Value of 83% improvement, G. Canyon (182)	\$82.7	0.95
Value of 376% improvement, G. Canyon (129)	\$128	2.2
Household income, \$1,000 (182)	\$28.1	0.10
Years of school attended (182)	14.1	0.17
Age in years (182)	41.0	0.12
Gender, male equals 0, female equals 1 (182)	0.637	0.0038
1981 Sample		
Value of 100% improvement in Chicago (71)	\$179	4.3
Value of 223% improvement in Chicago (71)	\$213	4.4
Value of 100% improvement in Chicago, and a 83% improvement in G. Canyon (71)	\$190	4.4
Household income, \$1,000 (71)	\$25.8	0.29
Years of school attended (71)	13.9	0.044
Age in years (71)	43.2	0.23
Gender, male equals 0, female equals 1 (71)	0.183	0.0056

^a Schooling, gender, and age statistics are for the respondent. The number of individuals in the sample is given in parentheses after each variable name.

^b In the 1980 sample, 129 respondents valued all three air quality improvements; 53 respondents valued only the 83% improvement.

quality range from \$58.8 to \$128 per year. The independent valuations of Chicago visibility range from \$179 to \$213 per year. The independent valuation of the 83% improvement in Grand Canyon visibility is \$82.7 per year. The same 83% change in Grand Canyon visibility is valued at only \$190 minus \$179 or \$11 per year when valued after a 100% change in Chicago visibility. The decrease in values due to the valuation sequence suggests strong substitution effects across regional air quality conditions.

Statistical tests for substitution were carried out using estimates of the quadratic and semilog bid functions. Table 2 reports the estimated equations. Variables used in the estimation were formulated in a manner consistent with equations (9) and (10). The dependent variable in all cases is the respondent's elicited valuation of the

described changes in visual air quality. The independent variables in quadratic form were entered in the manner described following equation (9). The independent variables in the semilog form were entered in the manner discussed following equation (10).

The independent variables are described in the first column of table 2. The first two variables listed, Grand Canyon Air Quality (*GCAQ*) and Chicago Air Quality (*CAQ*), measure the changes in air quality brought about by policy. The initial air quality level was normalized to one and the post-policy level was measured as a proportional change (e.g., 1.83, 2.0). Coefficients across the first two rows of table 2 measure elements of β for the quadratic and elements of γ for the semilog form.

The third, fourth, and fifth rows in table 2 give coefficients for the *B* and *D* matrices of, respectively, the quadratic and semilog functions. Rows three and four give the estimated coefficients for the diagonal terms. The fifth row states the off-diagonal terms. The remaining eight rows list the estimated elements of the *A* and *C* matrices that represent the interaction of air quality and socioeconomic characteristics.

Initial coefficient estimates were obtained using ordinary least squares (OLS). These are given in the second and fourth columns of table 2. Each of the air quality coefficients carries an intuitively consistent sign for environmental improvements that are amenities and substitutes. As air quality increases, a respondent's valuation increases but at a decreasing rate. The negative cross-product terms (fifth row) indicate substitution effects in the valuation of the two air quality variables.

Contingent valuation studies have generally ignored the inefficiency of OLS estimates when multiple valuations are obtained from the same individual respondents. An alternative is to specify the error structure as an error component model. An error component model allows for a nonzero covariance between a respondent's bids. With this model, the error terms in equations (9) and (10) are decomposed as $\epsilon_{rn} = \mu_r + \zeta_{rn}$ in the quadratic and $v_{rn} = \omega_r + \xi_{rn}$ in the semilog where $E(\mu_r^2) = \sigma_\mu^2$, $E(\omega_r^2) = \sigma_\omega^2$, and both ζ_{rn} and ξ_{rn} are uncorrelated across *r* and *n*. Generalized least squares (GLS) yields efficient estimates for the error component model (Judge et al. 1985, pps. 521–25).

The GLS estimates used for hypothesis testing are given in the third and fifth columns of table 2. The GLS data are identical to those used in the OLS equations. It is notable that the two

Table 2. Estimated Quadratic and Translog Value Functions

Variable ^a or Statistic	Quadratic ^{b,d}		Semilogarithmic ^{c,d}	
	OLS	GLS	OLS	GLS
Grand Canyon air quality (<i>GCAQ</i>)	132** (26.9)	121** (21.7)	181** (45.7)	177** (34.0)
Chicago air quality (<i>CAQ</i>)	234** (49.9)	238** (38.0)	345** (100)	349** (68.0)
(<i>GCAQ</i>) ²	-30.9** (10.3)	-28.3** (6.74)	-65.8* (38.8)	-63.4** (23.2)
(<i>CAQ</i>) ²	-66.1** (22.5)	-66.1** (13.8)	-160* (90.0)	-160** (49.4)
<i>GCAQ</i> × <i>CAQ</i>	-46.8** (25.6)	-42.6** (13.9)	-87.2** (50.9)	-85.2** (28.2)
<i>GCAQ</i> × household income (\$1,000)	0.706 (0.547)	0.491 (0.372)	29.9 (23.9)	21.4 (16.9)
<i>GCAQ</i> × years in school	-0.902 (3.27)	0.814 (2.26)	-21.3 (86.5)	33.1 (64.0)
<i>GCAQ</i> × age in years	-1.24** (0.454)	-0.910** (0.317)	-92.9** (33.0)	-72.2** (25.5)
<i>GCAQ</i> × gender	-12.2 (14.5)	-7.82 (11.2)	-21.1 (27.6)	-16.1 (21.5)
<i>CAQ</i> × household income (\$1,000)	-1.41** (0.541)	-0.736 (0.502)	-58.8** (21.7)	-41.6* (23.5)
<i>CAQ</i> × years in school	15.0** (3.59)	6.63** (3.33)	339** (87.3)	224** (94.7)
<i>CAQ</i> × age in years	1.48** (0.646)	0.392 (0.601)	137** (49.0)	68.6 (53.5)
<i>CAQ</i> × gender	-26.0 (27.2)	-19.8 (25.3)	-49.3 (47.5)	-44.5 (51.9)
Root mean square error	233	112	233	111
<i>R</i> -squared	0.28	0.22	0.28	0.23
<i>F</i> -value	19.5	14.0	19.5	14.3

^a Variables are described in table 1.

^b Variables in the quadratic were formulated as in equation (9).

^c Variables in the semilogarithmic were formulated as in equation (10) except for the dummy variable gender. The dummy variable was entered without transforming it by a logarithm.

^d Columns labeled OLS were estimated using ordinary least squares. Columns headed GLS were estimated using the error components model described in Judge et al. (1985, pp. 521–25). A double pound sign (+) indicates that a coefficient is significantly different from zero at the 95% level for a one way test. A single (double) asterisk indicates that a coefficient is significantly different from zero at the 90% (95%) level for a two-way test. Standard errors are in parentheses. Estimates have 640 degrees of freedom.

sets of coefficients are very similar. As with the OLS estimates, air quality coefficients indicate that values increase at a decreasing rate with increases in air quality. The cross-quality interaction terms again suggest that regional air quality conditions are substitutes in valuation. Finally, consistent with the theoretical efficiency of the GLS estimator, the GLS estimated root mean square error is about 50% smaller than that of OLS for both the quadratic and semilog models.

Four sets of hypothesis tests were carried out with the GLS equations. First, with the derived functional forms, theory implies a bid function (1) without an intercept and (2) where the socioeconomic variables enter only as interaction terms with the air quality variables. For the quadratic, an *F*-test failed to reject the joint hypothesis of α_0 and α_s equal to zero (*F*-value of 1.0; 5 and 635 degrees of freedom). An *F*-test

also failed to reject the hypothesis of ρ_0 and ρ_s equal to zero in the semilog model (*F*-value of 0.53; 5 and 635 degrees of freedom).

Second, the inequality restrictions imposed by concavity are testable using the likelihood ratio method. This method compares a set of restricted estimates with the unrestricted estimates given in table 2. Examination of table 2, however, shows that the concavity restrictions are nonbinding. The diagonal elements of *B* and *D* are strictly negative and $|B|$ and $|D|$ are strictly positive. In this nonbinding case, restricted least squares estimates are identical to the unrestricted estimates (Judge et al. 1988). The results, therefore, fail to reject a concave bid function.¹²

¹² Diewert and Wales, and Geweke present alternative estimators for the general case where concavity restrictions may be binding.

Third, the significance of the substitution effect was examined. The coefficient of $GCAQ \times CAQ$ is negative and different from zero at any conventional level of significance in both the quadratic and semilog equations. These results confirm the predictions of theorem 1 and reject both independence and complementarity. Regional air quality conditions are substitutes in valuation.

Finally, a J-test (MacKinnon) was used to test the quadratic and semilog specifications. With the J-test, the quadratic and semilog specifications are tested against each other. To test the quadratic, the predicted values from the estimated semilog equation are entered as an additional explanatory variable in the quadratic equation and the equation is reestimated. If the coefficient on the predicted bid from the semilog equation is significantly different from zero, the predicted value contributes to the quadratic's explanatory power and the quadratic specification is rejected as incomplete. The predicted bid from the estimated quadratic equation is used to test the semilog in a similar manner.

J-tests were conducted for both the quadratic and the semilog. In the quadratic equation, the estimated coefficient for the predicted hc was 1.09 and had a t -value of 1.9 with 639 degrees of freedom. This result rejects the quadratic as a valid specification at the 90% significance level. In the semilog equation, the coefficient for the predicted hc was 0.0668 and had a t -value of 0.11 with 639 degrees of freedom. This result fails to reject the semilog specification at any conventional significance level. The semilog is therefore the better fitting specification for these data.

Concluding Comments

Environmental values are contextual—they are conditioned on the presence or absence of multidimensional resource flows. A conventional benefit cost design ignores the contextual nature of value data. Where resources are substitutes in valuation, a conventional benefit cost estimate exceeds the valid measure of net benefits. Where complementarity is present, a conventional benefit cost estimate understates the valid valuation. An accurate resource policy valuation requires a valid valuation design.

A valid valuation design has two characteristics. First, the aggregate valuation of a policy change is unique. For any given policy, there is only one valid valuation. Second, a theoretically

consistent disaggregated valuation can only be obtained by valuing policy impacts along a path of valuation. A valid path accounts for both the change in a specific policy component as well as the overall policy change.

The valid benefit cost design reduces to the conventional approach only if resource services are independent in valuation. Theorem 1 shows that such independence is unlikely. For a small number of impacts, substitution effects are expected for environmental services that are separated in consumption by space or uncertainty. Complementarity is possible if environmental services are tied together by a specific consumption activity.

The restrictions of a valid design apply to both contingent valuation and demand-based methods such as travel cost and hedonic analysis. Well-known procedures exist for estimating substitution and complementary effects in demand systems. Because the problem of interaction is new to the analysis of contingent valuation data, this paper has examined two different ways to implement a theoretically consistent contingent valuation. First, if the policy alternatives are predetermined, contingent valuation can be used to holistically evaluate benefits in a one-step, aggregate valuation.

Second, a sample of contingent value data may be used to estimate a multidimensional bid function. The estimation problem is similar to estimating a multiproduct cost function—the only requirements are a set of admissible functional forms and statistical design that elicits contingent values for a matrix of proposed environmental changes. The estimated bid function may be used to evaluate any combination of policy impacts within the range of a sample data set.

The econometric approach was empirically demonstrated using contingent valuation data for regional air quality. Quadratic and semilogarithmic bid functions were derived and estimated. Both functions conformed to the requirements of theory. Statistical specification tests indicated that the semilogarithmic function provided a better fit to the data used in the case study. Consistent with theoretical predictions, regional environmental conditions were shown to be strict substitutes in valuation.

These results should encourage further research to improve understanding of interaction effects in valuation. A range of alternative functional forms exist (Griffin, Montgomery, and Rister) that may be applied to the contingent valuation problem and tested for robustness and consistency with theory (Thompson).

In addition, the extent and strength of substitution should be examined in alternative valuation contexts. Intraregional packages of amenities offer cases of potential complementarity. These cases should be investigated to determine whether complementarity in activity production outweighs the substitution effects inherent in constrained valuation.

[Received April 1989; final revision received May 1990.]

References

- Baumol, William J., John C. Panzar, and Robert D. Willig. *Contestable Markets and the Theory of Industry Structure*. New York: Harcourt, Brace, Jovanovich, 1982.
- Bockstael, Nancy E., and Catherine L. Kling. "Valuing Environmental Quality: Weak Complementarity with Sets of Goods." *Amer. J. Agr. Econ.* 70(1988):654–62.
- Bradford, D. F. "Benefit-Cost Analysis and the Demand for Public Goods." *Kyklos* 23(1970):775–91.
- Braeutigam, Ronald, and Roger G. Noll. "The Regulation of Surface Freight Transportation: The Welfare Effects Revisited." *Rev. Econ. and Statist.* 66(1984):80–87.
- Brookshire, David S., Alan Randall, and John R. Stoll. "Valuing Increments and Decrements in Natural Resource Flows." *Amer. J. Agr. Econ.* 62(1980):478–88.
- Burt, Oscar, and Durward Brewer. "Estimation of Net Social Benefits from Outdoor Recreation." *Econometrica* 39(1971):813–28.
- Caves, Douglas W., and Laurits R. Christensen. "Global Properties of Flexible Functional Forms." *Amer. Econ. Rev.* 70(1980):422–32.
- Deaton, Angus, and John Muellbauer. *Economics and Consumer Behavior*. London: Cambridge, 1980.
- Diewert, W. E., and T. J. Wales. "Flexible Functional Forms and Global Curvature Conditions." *Econometrica* 55(1987):43–68.
- Fischhoff, Baruch, and Lita Furby. "Measuring Values: A Conceptual Framework for Interpreting Transactions with Special Reference to Contingent Valuation of Visibility." *J. Risk and Uncertainty* 1(1988):147–84.
- Gallant, A. Ronald. "The Fourier Flexible Form." *Amer. J. Agr. Econ.* 66(1984):204–8.
- Geweke, J. "Exact Inference in the Inequality Constrained Normal Linear Regression Model." *J. Appl. Econometrics* 1(1986):127–41.
- Green, H.A. John. *Aggregation in Economic Analysis*. Princeton, NJ: Princeton University Press, 1964.
- Griffin, Ronald C., John Montgomery, and M. Edwards Rister. "Selecting Functional Form in Production Function Analysis." *West. J. Agr. Econ.* 12(1987):216–27.
- Hoehn, John P. "The Benefit Cost Evaluation of Multipart Public Policy: A Theoretical Framework and Critique of Estimation Methods." Ph.D. thesis, University of Kentucky, 1983.
- Hoehn, John P., and Alan Randall. "Too Many Proposals Pass the Benefit Cost Test." *Amer. Econ. Rev.* June (1989).
- Hof, John G., Robert D. Lee, A. Allen Dyer, and Brian M. Kent. "An Analysis of Joint Costs in a Managed Forest Ecosystem." *J. Environ. Econ. and Manage.* 12(1985):338–52.
- Judge, George G., William E. Griffiths, R. Carter Hill, Helmut Lütkepohl, and Tsoung-Chao Lee. *Introduction to the Theory and Practice of Econometrics*. New York: John Wiley & Sons, 1988.
- . *The Theory and Practice of Econometrics*. New York: John Wiley & Sons, 1985.
- Kahneman, Daniel, and John Knetsch. "Valuing Public Goods: The Purchase of Moral Satisfaction." University of California work. pap., 1990.
- Knetsch, Jack. "Displaced Facilities and Benefit Calculations." *Land Econ.* 53(1977):123–29.
- MacKinnon, James G. "Model Specification Tests Against Non-Nested Alternatives." *Econometric Rev.* 2(1983):85–110.
- Majid, I., J. A. Sinden, and Alan Randall. "Benefit Evaluation of Increments to Existing Systems of Public Facilities." *Land Econ.* 59(1983):377–92.
- Mäler, Karl-Göran. *Environmental Economics: A Theoretical Inquiry*. Baltimore MD: Johns Hopkins University Press, 1974.
- Mitchell, Robert Cameron, and Richard T. Carson. *Using Surveys to Value Public Goods: The Contingent Valuation Method*. Washington DC: Resources for the Future, 1989.
- Randall, Alan, and John R. Stoll. "Existence Value in a Total Valuation Framework." *Managing Air Quality and Scenic Resources at National Parks and Wilderness Areas*, ed Robert D. Rowe and Lauraine G. Chestnut. Boulder CO: Westview Press, 1983.
- Rosenthal, Donald. "The Necessity for Substitute Prices in Recreational Demand Analyses." *Amer. J. Agr. Econ.* 69(1987):828–37.
- Thompson, Gary D. "Choice of Flexible Functional Forms: Review and Appraisal." *West. J. Agr. Econ.* 13(1988):169–83.
- Tolley, George, and Robert Fabian. *The Economic Value of Visibility*. Mount Pleasant MI: Blackstone, 1988.

Tax Versus Quota Regulation: A Stochastic Model of the Fishery

Robert A. Androkovich and Kenneth R. Stollery

In a deterministic setting, it has been demonstrated that taxes and quotas can be equivalent regulatory instruments and that both can in principle induce an optimal allocation of resources in markets where externalities are problematic. In this paper we consider growth, harvest, and demand uncertainty within the context of Canada's Pacific halibut fishery and find that, while taxes remain the preferred instrument, the welfare losses arising from the use of individual boat quotas are minor.

Key words: externality, quota, regulation, taxation.

There is wide consensus on the need for regulation of industries such as fisheries where significant externalities exist and on the standard remedies to be applied. For example, in situations where the regulatory agency has full information and administration costs are equal, it is well known that per-boat quotas, taxes on the catch, and even per-boat effort controls are equivalent regulatory methods in theory. In practice, however, there is uncertainty about demand conditions, the size of the stock, and the productivity of the fleet; this situation not only creates great difficulty in actual resource management but prevents the alternative methods of regulation from being theoretically equivalent. In the work derived from Weitzman's general discussion of regulation under uncertainty, the ranking of taxes and quotas depends, for instance, upon the relative slopes of the marginal benefit and marginal social cost curves and whether the fish stock is observable without error (see Andersen, Anderson, Koenig 1984a).

This paper differs from the Weitzman approach with respect to the timing of decisions made by the regulatory authority and fishermen. Whereas in Weitzman's paper decisions are taken *ex post*, that is after the realization of the random variables, in the present model all economic agents take their decisions *ex ante*.¹ This approach is more realistic in that a boat owner

would typically have to decide whether to enter a particular fishery prior to knowing the price of fish, rather than after.² It turns out that this assumption guarantees that an appropriately set specific tax will be capable of inducing optimal behavior by the fishermen. This does not, of course, follow when decisions are made *ex post*.

Previous models introduce quotas that bind individual fishermen to produce a fixed amount. However, because fishing is inherently risky for participants as well as regulators, there is no guarantee that the quota will be met. Instead, this analysis follows Clark (1985b) by introducing per-boat quotas which truncate the distribution of catch, with catch above the quota returned to the ocean. As is well known, the mortality rate for returned fish is quite high. The importance of a potential divergence between the quantity of fish consumed and killed is obvious in that the future fish stock would be less than would be predicted on the basis of the growth equation and the quota.

In order to illustrate the extended quota model, the welfare effects of taxes and quotas are compared in the context of the Pacific halibut fishery, one with a homogenous fishing technology and fairly reliable sources of data. This comparison is done by simulating a stochastic dynamic program with parameters obtained from studies of this fishery. While taxes in the model are first-best in a limited-information sense, the

Robert A. Androkovich and Kenneth R. Stollery are, respectively, an assistant professor and a professor in the Department of Economics, University of Waterloo, Ontario.

The authors are indebted to three anonymous referees for their suggestions.

¹ Anderson makes a similar assumption.

² We are abstracting here from possible sharecontracting arrangements with processors. In many fisheries, processors absorb some of the risk by contracting for a minimum price on delivery. We are indebted to an anonymous referee for this observation.

welfare loss from individual quotas is relatively small.

The outline of the paper is as follows: the second section outlines the fishery model and compares the competitive, optimal, taxation, and quota equilibria. The sources of the parameters required for simulation of the model then are discussed, preparatory to the simulation of the model in the subsequent section. The final part offers conclusions.

The Fishery Model

The assumed fishery production function is a modified Schaefer type, given by

$$(1) \quad Q(t) = N(t)g[X(t), \theta(t)]e(t),$$

where $Q(t)$ is aggregate harvest during period t , $e(t)$ is effort per boat, $N(t)$ is number of (identical) boats fishing, $X(t)$ is stock size, $g[X(t), \theta(t)]$ is catch per unit effort (*CPUE*), and $\theta(t)$ is a random variable with mean unity which introduces output uncertainty.

The particular *CPUE* function is given by $g[X(t), \theta(t)] = \theta(t)/[\omega - \sigma X(t)]$, where ω and σ are parameters. The specific and somewhat unusual form of this function is required to make the subsequent analysis tractable. Quadratic benefit and per-boat cost functions are also required for a closed-form solution to the subsequent dynamic programming problem. Thus, an aggregate catch of $Q(t)$ has associated with it total benefits to consumers of

$$(2) \quad B[Q(t)] = [b_0 + \beta(t)]Q(t) - b_1 Q(t)^2/2,$$

and per-boat total costs of

$$(3) \quad C[e(t)] = c_0 + c_1 e(t) + c_2 e(t)^2/2;$$

with parameters $b_0, b_1, c_0 > 0$, one or more of c_1 and $c_2 > 0$, and $\beta(t)$ a random variable (with mean zero) which accounts for demand uncertainty.³ These functions imply a linear demand relationship for fish landings of the form $P = b_0 - b_1 Q$ (coinciding with that estimated by Cook for the halibut fishery) and a linear marginal cost curve for fishery effort. Given the paucity of data on effort costs, estimation of a nonlinear marginal cost function, even if theoretically justifiable, would not be possible with the available information. Following Koenig (1984a,b), we

also employ a linear growth relationship, approximating the growth of the fish stock by

$$(4) \quad X(t+1) = \phi_0 + \phi_1(X(t) - Q(t)) + \varepsilon(t);^4$$

where $\phi_0 \geq 0$, $\phi_1 > 0$, and $\varepsilon(t)$ is a random variable with mean zero.

The particular forms chosen for the benefit, cost, and growth relationships guarantee that the dynamic programming technique permits the conversion of the problem into one involving only current period variables. This, in turn, simplifies the comparison of the alternative regulatory regimes with the competitive case.

The regulatory authority's goal is the maximization of the present discounted value of the fishery's expected net benefits to society: equivalently, the sum of consumer surplus and economic rent. This is given by

$$(5) \quad V_0 = E_0 \sum_{t=0}^{\infty} \rho^t [B[Q(t)] - N(t)C[e(t)]];^5$$

where $0 < \rho < 1$ is the discount factor.⁵ Applying standard dynamic programming techniques,⁶ let

$$(6) \quad V_j[X(j)] = \text{Max}_{e(j)} E_j \sum_{t=j}^{\infty} \rho^t [B[Q(t)] - N(t)C[e(t)]].$$

Then,

$$(7) \quad V_t[X(t)] = \text{Max}_{[e(t), N(t)]} E_t [B[Q(t)] - N(t)C[e(t)] + \rho V_{t+1}[X(t+1)]],$$

and it is straightforward, though tedious, to show that the solution to this recursive relationship is given by

$$(8) \quad V_t[X(t)] = a_0(t) + a_1(t)X(t) + a_2(t)X(t)^2/2,$$

where the a_i also follow recursive relationships given by⁷

⁴ The long-lived nature of halibut makes it a multicohort fishery. If age-selectivity were allowed when setting catch quotas, the optimal quota could be determined by employing a Beverton-Holt population model in terms of numbers of fish, combined with a function such as Von Bertalanffy's describing the growth in weight of a single fish with age. This refinement was not considered in this paper because finding a closed-form solution to the dynamic program restricts the population growth function to a single linear form. A population model in terms of total biomass is equivalent to the disaggregated model in a steady-state, where the age distribution is constant.

⁵ $E_Z\{X\}$ is the expectation of $\{X\}$ at time Z .

⁶ See Chow, especially chapters 7 and 8.

⁷ In obtaining relationships (8)–(11), we assume an optimal policy is followed in all future time periods and that the three random variables are uncorrelated both inter- and intratemporally.

³ The demand uncertainty takes the form of a linear shift in the demand relationship.

$$(9) \quad a_0(t-1) = \rho[a_0(t) + a_1(t)\phi_0] + \rho a_2(t) \frac{[\phi_0^2 + \text{Var}(\varepsilon)]/2 + \chi^2/2\xi,}{\xi},$$

$$(10) \quad a_1(t-1) = \rho\phi_1[a_1(t) + a_2(t)\phi_0] + \chi[\sigma\Omega - \rho a_2(t)\phi_1^2]/\xi, \text{ and}$$

$$(11) \quad a_2(t-1) = \rho a_2(t)\phi_1^2 + [\sigma\Omega - \rho a_2(t)\phi_1^2]^2/\xi;$$

where $\Omega \equiv [c_1 + \sqrt{2c_0c_2}]$, $\xi \equiv E(\theta^2)[b_1 - \rho a_2(t)\phi_1^2]$, and $\chi \equiv [b_0 - \rho a_1(t)\phi_1 - \rho a_2(t)\phi_0\phi_1 - \omega\Omega]$.

It follows from (4), (5), and (8) that in the current period, the regulatory authority should choose both the number of boats and effort per boat so as to maximize the extended net benefit function⁸

$$(12) \quad ENSB = E\{B[Q] - NC[e] + \rho[a_0 + a_1[\phi_0 + \phi_1(X - Q) + \varepsilon] + a_2[\phi_0 + \phi_1(X - Q) + \varepsilon]^2/2]\}.$$

The first-order conditions for an interior maximum are given by

$$(13) \quad E\{[b_0 + \beta - b_1Q]g[X, \theta] - C'[e] - \rho\phi_1E\{[a_1 + a_2[\phi_0 + \phi_1(X - Q) + \varepsilon]]g[X, \theta]\} = 0, \text{ and}$$

$$(14) \quad E\{[b_0 + \beta - b_1Q]g[X, \theta]e\} - C[e] - \rho\phi_1E\{[a_1 + a_2[\phi_0 + \phi_1(X - Q) + \varepsilon]]g[X, \theta]e\} = 0.$$

The above characterization of the optimal harvest assumes the government has total control over the fishery. The equilibrium also represents a limited information optimum in that the government does not have knowledge of the realized values of the random variables when decisions are made. This situation represents a benchmark against which to compare more realistic cases in which the government either has no control over the fishery or controls it indirectly through the use of taxes or quotas.

A competitive situation can be modelled by setting $a_0 = a_1 = a_2 = 0$ and maximizing (12) with respect to e , with entry driving expected fishery rents to zero.⁹ This yields

$$(15) \quad E\{[b_0 + \beta - b_1Q]g[X, \theta]\} = C'[e] \text{ and}$$

$$(16) \quad E\{[b_0 + \beta - b_1Q]g[X, \theta]e\} = C[e]$$

as the equilibrium conditions. Although yielding the optimum level of per-boat effort, the usual stock externality represented by the absence of the a_i terms means a competitive fishery will generally lead to overutilization of the stock. An issue of primary concern is the extent to which either a specific tax or quota will mitigate the externality problem in a stochastic setting.¹⁰

With a specific tax (τ), effort will be expended and entry will occur until

$$(17) \quad E\{[b_0 + \beta - b_1Q - \tau]g[X, \theta]\} = C'[e], \text{ and}$$

$$(18) \quad E\{[b_0 + \beta - b_1Q - \tau]g[X, \theta]e\} = C[e].$$

Acting as a principal with the fishermen as agents, the government will then

$$(19) \quad \text{Max}_{[\tau]} E\{B[Q] - NC[e] + \rho[a_0 + a_1[\phi_0 + \phi_1(X - Q) + \varepsilon] + a_2[\phi_0 + \phi_1(X - Q) + \varepsilon]^2/2]\}$$

subject to (17) and (18); in the process all fishery rents becoming tax revenues.¹¹ The resulting first-order condition can be written in the form

$$(20) \quad E\{[b_0 + \beta - b_1Q]g[X, \theta]e\} - C[e] - \rho\phi_1E\{[a_1 + a_2[\phi_0 + \phi_1(X - Q) + \varepsilon]]g[X, \theta]e\} = 0.$$

It follows immediately, from the identity of equations (14) and (20), that if the regulatory authority chooses a specific tax which satisfies $\tau = \rho\phi_1E\{[a_1 + a_2[\phi_0 + \phi_1(X - Q)]]\}$, the number of boats in the fishery, the effort per boat, and the level of expected net social benefits will be identical to their values in the limited-information optimal situation. Thus, a specific tax can, in principle, induce a competitive fishery to harvest optimally in a stochastic as well as a deterministic environment.

It is of interest to compare the optimality of taxation in our model with the conclusions derived by Koenig (1984a,b 1985), who finds tax-

⁸ Because the problem has been converted into one involving only current period variables, time references are dropped in the formulas which follow. a_0 , a_1 , and a_2 represent shadow prices in (12), in that they account for the effects of current variables on future net benefits.

⁹ This formulation assumes fishery firms are risk-neutral. The justification for this assumption derives from the finding by Andersen that risk aversion by fishermen under stock or price uncertainty reduces both effort and expected catch, possibly even to sub-optimal levels. Because overfishing motivates regulatory intervention in the first place, by assumption the degree of risk aversion on the part of fishing firms must be comparatively low.

¹⁰ In a deterministic environment a specific tax can, in principle, eliminate the externality problem if the tax is set equal to the shadow value of the stock.

¹¹ The costs associated with administering differing regulatory methods are not considered. Different costs of administering taxes and quotas would affect the losses in expected net social benefits associated with their respective usage. For instance, if a quota regime were less costly to manage than a tax regime, then the results summarized in table 1 might be reversed, with a quota dominant.

ation a suboptimal instrument. The difference arises from the divergent treatments of uncertainty in our paper and Koenig's. He assumes that while output decisions are made with full information—after the realization of the random variables—the rates of taxation are set *ex ante*. This asymmetry of information, in turn, guarantees that taxes are suboptimal in a full-information sense.¹² We assume instead that decisions are made by both the regulatory authority and the boat owners prior to the realization of the random variables; that is, the size of the fleet will be nonrandom *ex ante*, and the decisions regarding the rate of taxation and whether to fish must both be made with incomplete information. This symmetry of information within our model guarantees that a specific tax will be first-best given the limited information.

erwise, the future stock will be less than would be predicted from the growth equation and the quota, and the fishery's expected net benefits to society would be inaccurate if this issue were to be ignored. To illustrate, define m as the mortality rate for returned fish, and

$$S = \begin{cases} 1 & \text{if } \theta \leq \theta^* \\ 0 & \text{if } \theta > \theta^* \end{cases}.$$

Then, actual (per-boat) mortality is

$$\bar{q} = Sg[X, \theta]e + (1 - S)\{q + m(g[X, \theta]e - q)\}.$$

It follows that if individual boat quotas are used, the level of the fisheries expected net benefits to society is given by

$$(21) \quad ENSB[q, N, e] = E \left[\int_0^{\theta^*} B[Q]f(\theta)d\theta + \int_{\theta^*}^{\infty} B[Nq]f(\theta)d\theta - NC[e] + \rho \int_0^{\infty} \left[a_0 + a_1[\phi_0 + \phi_1(X - N\bar{q}) + \varepsilon] + a_2[\phi_0 + \phi_1(X - N\bar{q}) + \varepsilon]^2/2 \right] f(\theta)d\theta \right];$$

The manner in which we model risk also implies that when individual boat quotas are used as the method of regulation, they change the harvest function by truncating the catch distribution. The harvest function becomes

$$Q = \begin{cases} Ng[X, \theta]e & \text{if } \theta \leq \theta^* \\ Nq & \text{if } \theta > \theta^* \end{cases};$$

where q is the per boat quota, and $\theta^* = [\omega - \sigma X]q/e$ is the value of θ associated with a catch of q , stock level X , and per-boat effort e .

Of course, the possibility of a catch in excess of the quota now arises. We assume that any such excess is immediately returned to the ocean. It is known that the mortality rate for returned fish is quite high; thus in modeling a quota under this type of uncertainty it is important to account for a potential divergence between the quantities of fish consumed and killed.¹³ Oth-

where $f(\theta)$ is the probability density function for the random variable θ , and the expectations operator refers to the random variables β and ε . Acting once again as a principal, the regulatory authority will then

$$(22) \quad \max_{[q, N, \hat{e}]} ENSB[q, N, \hat{e}] \text{ subject to}$$

$$\hat{e} = \operatorname{argmax}_{[e]} \int_0^{\infty} [b_0 - b_1 Q]g[X, \theta]ef(\theta)d\theta - C[e].$$

The constraint, equation (22), implies that each boat owner applies effort to the point where his expected marginal value equals his marginal cost.¹⁴ The regulatory authority, being aware of the boat owner's decision-making processes, accounts for this in setting the quota and the fleet size.¹⁵

¹² Within the context of the model specified, this situation would involve boats entering or leaving the fishery depending upon the realized values of the random variables.

¹³ Myhre estimates the mortality of released halibut at 32%; this figure was used in the simulation analysis to follow. Results were not highly sensitive to the value of this parameter.

¹⁴ In characterizing \hat{e} , the boat owners are assumed to take the price and their individual quotas as given.

¹⁵ A situation in which an industry-wide quota is established, in conjunction with control over the size of the fleet, is formally equivalent to the one which we have modeled since the boats in the fleet are homogenous both in their use of inputs and in the mortality of fish returned to the sea. The approach we follow in solving the quota problem is analogous to that which would be followed in (numerically) solving a boundary value problem.

Sources of Model Parameters

In order to gain insight into the welfare costs associated with the implementation of quotas, a specific fishery is considered. Parameters are required for the demand and cost of effort functions and the fishery biological relationship. The Pacific halibut fishery was chosen to illustrate the model solution because it has a relatively simple (longline) fishing technology, as well as a long regulatory history giving rise to reliable data. For these reasons this industry has been quite extensively studied, and parameter estimates from this previous research are employed.

Cook estimated a linear (inverse) demand relationship for regulatory area 2 (mainly the Canadian sector). With the catch measured in tons, the price in 1961 dollars, and standard errors in parentheses, the estimated function was

$$P = 720.0 + 2220.0P_s - 1.02SL - 0.04Q$$

$$(171.4) \quad (325.4) \quad (0.54) \quad (0.0098)$$

$$\bar{R}^2 = 0.77, D-W = 1.62,$$

Method: ordinary least squares

where P_s is the price of salmon (a major substitute) and SL is the length of the fishing season (affecting the catch quality).¹⁶ For the simulation, this equation was collapsed by substituting their average 1968–80 values, giving

$$P = 1200.0 - 0.04Q$$

$$(603.3) \quad (0.0098)$$

Because this represents the marginal social benefit from harvesting area 2, the estimates of b_0 and b_1 are 1200.0 and 0.04, respectively, with the variance of β derived from the b_0 coefficient's standard error, becoming $\text{var}(\beta) = 1.75 \times 10^4$.

The parameters for the per-boat cost of effort function were established using Catch Summary Reports from 1975–80 on the Canadian Pacific halibut fleet made available by the Canadian Department of Fisheries and Oceans (Stollery).

Halibut vessels were defined as those taking 80% or more halibut. The reports showed the fleet to be relatively homogenous, for of ninety-seven boats in 1979, eight-three were between 30 feet and 60 feet in length, and these represented 89% of the total catch. They also showed an average capital value of \$491,293 per boat in 1961 dollars. Employing a real discount rate of 5%, 4.4% capital depreciation, and insurance costs of 3.8% of vessel value, yearly fixed costs per boat are estimated at 13.2% of capital value, or $c_0 = \$64,895.00$. The values for depreciation and insurance were not available in the Catch Summary Reports, but were taken from "Basic Economic Indicators, Halibut," a 1968 survey of the U.S. Pacific Halibut fleet by the U.S. Bureau of Commercial Fisheries, Department of the Interior, 1970. Prior to 1979–80, when a limited-entry licensing scheme was first imposed, regulation in the halibut fishery consisted only of controls on season length (Cook and Copes). Thus, the period before 1980 represents open-access, and the model is calibrated to a competitive benchmark. Cook and Copes' data imply an average effort per boat of 1357.5 skate-soaks over the 1968–80 period.¹⁷ Equating the marginal and average costs of effort at this value gives $c_1 = -\$55.55$ and $c_2 = \$0.07$ for the remaining cost function parameters.

The catch per unit effort ($CPUE$) function is specified in the model as $g[X, \theta] = \theta/[\omega - \sigma X]$.¹⁸ The parameters ω and σ are obtained by taking a first-order Taylor-series approximation of expected industry $CPUE$, defined as $\bar{g}[X] = E\{g[X, \theta]\} = [\omega - \sigma X]^{-1}$, at \bar{X} (the mean stock level), then constraining this approximation to the traditional Schaefer function $g_0 X$ at \bar{X} . Equating $\bar{g}[X] = \bar{g}[\bar{X}] + \sigma \bar{g}[\bar{X}]^2 (X - \bar{X})$ with $g_0 X$ implies $\sigma = (\bar{X} \bar{g}[\bar{X}])^{-1} = \omega/2\bar{X}$ and $4\sigma/\omega^2 = g_0$. Employing data prior to limited-entry licensing in 1980, Cook calculated the mean stock level implied by combining the Schaefer production function and the Verhulst (logistic) population growth equation at bioeconomic equilibrium, giving $\bar{X} = 29,334$ tons. This gives $\bar{g}[\bar{X}] = 0.04328$ tons per skate-soak and $g_0 = 0.12 \times 10^{-5}$, implying $\sigma = 7.875 \times 10^{-4}$ and $\omega = 46.216$ for the $CPUE$ parameters.

Because the dispersion of the catch each year

¹⁶ The salmon price was exogenous from the standpoint of the regulators, but they were in fact controlling the season length during the period prior to the imposition of limited-entry licensing in the 1980s. We have not modeled season controls explicitly as a regulatory tool; but, from the standpoint of controlling fishery effort, an indirect benefit of such control in our model would be a reduced variance associated with the estimation of social benefits. Season length controls may also mean that SL and Q were correlated over the estimation period, possibly introducing multicollinearity. As long as the equation errors are independently distributed this will not bias the coefficients, but the coefficient standard errors (and thus the estimate of β 's variance) could have been affected by this statistical problem.

¹⁷ In the halibut fishery effort is measured in "skate-soaks" per time period. A skate is a (standard) length of fishing line which is left in the water (soaked) for a set amount of time.

¹⁸ This particular form for the $g[X, \theta]$ function was required to make the model tractable. It is an approximation to $g[X, \theta] = g_0 \theta X$, i.e., the traditional Schaefer harvest function.

around its mean value is not known, no information on the variance of θ is available from this method. In 1979, the year prior to the imposition of limited-entry licensing, ninety-seven boats were in the fishery. In order to calibrate the competitive equilibrium to this benchmark, appropriate values must be specified for any parameters lacking empirical verification. Only the variance of θ falls into this category. Thus, the specification of $\text{Var}(\theta) = 0.225$ ensures that the competitive equilibrium is realistic.¹⁹

The parameters ϕ_0 and ϕ_1 characterize the yearly growth of the fish stock [equation (4)]. This linear approximation to the traditional Verhulst logistic growth function is also necessary in order to solve the stochastic dynamic programming problem. Because stock size is not directly observable, (4) cannot be estimated directly; the only available proxy is industry-wide CPUE or $g[X]$.²⁰ For purposes of estimation, we assume $g[X] = g_0X$, which in conjunction with (4) yields

$$CPUE_t = \phi_0 g_0 + \phi_1 [CPUE_{t-1} - g_0 Q_{t-1}] + g_0 \epsilon.$$

This, of course, could be estimated using nonlinear least squares to obtain ϕ_0 , ϕ_1 , and g_0 . However, Cook estimated the traditional logistic growth function for the stock, providing information on g_0 .²¹ Therefore, g_0 is constrained to the estimate of 0.12×10^{-5} , obtaining $\phi_0 = 6259.63$ and $\phi_1 = 0.87216$. The variance of ϵ was determined as the regression error variance divided by g_0^2 , becoming $\text{Var}(\epsilon) = (7921.84)^2$.

Given the parameter values above, the coefficients a_0 , a_1 , and a_2 characterize the optimal $V(X)$ function, i.e., the expected value of the

fishery's net benefits to society. As was pointed out previously, these coefficients follow recursive relationships: functions $f_0[\cdot]$, $f_1[\cdot]$, and $f_2[\cdot]$ exist such that

$$a_0(t-1) = f_0[a_0(t), a_1(t), a_2(t); \text{parameters}],$$

$$a_1(t-1) = f_1[\dots], \text{ and}$$

$$a_2(t-1) = f_2[\dots].$$

The CHOW-YORKE fixed-point algorithm from the ACM algorithms distribution service (see Watson, Watson and Fenner) was used to determine their convergent values; $a_0^* = 10006026.8283$, $a_1^* = -505.4409$, and $a_2^* = 0.0247$. These values were substituted into the various first-order conditions, allowing us to determine, for the alternative cases, the optimal number of boats, effort, and the maximum expected net social benefits. Where required, the BRENTM subroutine (also available from ACM) was used to solve the systems of nonlinear equations, and all integration was completed using the subroutine RKF45 (Forsythe, Malcolm, and Moler).

Simulation Results

Recall that in a first-best situation, the government has complete control of the fishery and chooses the level of effort per boat and the number of boats so as to maximize the fishery's expected net benefits to society. Solving the dynamic program with the parameter estimates, effort is 1,357.5 skate-soaks per year, approximately seventy-one boats are in the fishery, and expected output is 4,155.3 tons with expected net social benefits of \$5,819,967.34, as shown in the first column of table 1.

In comparing the optimal fishery with the competitive case, the fleet under competition is larger (by an estimated 37%) and the expected net social benefits of the fishery are correspondingly reduced. The effort levels in the first-best and competitive cases are equal because the form that has been adopted for the CPUE function ensures that, in both cases, effort is applied to the point of minimum average cost of effort. Expected net benefits are positive even in the competitive case because the price will equal marginal benefits and will be exceeded by average benefits.

We have indicated how a specific tax can induce optimal behavior in this model. With the parameters the optimal tax (τ) is \$74.35 per ton,

¹⁹ For the simulation experiment, we assume θ is distributed according to a gamma density function, i.e.,

$$f(\theta) = \lambda^\tau \theta^{\tau-1} e^{-\lambda\theta} / \Gamma(\tau), E[\theta] = \tau/\lambda, \text{ and } \text{Var}[\theta] = \tau/\lambda^2;$$

where $\Gamma(\tau) = \int_0^\infty x^{\tau-1} e^{-x} dx$ and $r = \lambda = 4.89$.

²⁰ CPUE is a nonlinear function of X according to the yield function, but we regard this as a necessary approximation. If $g[X]$ were a correct representation of the yield-stock relationship in the fishery, the stock growth function parameters could be estimated by inverting $g[X]$ in order to obtain effort per unit catch (EPUC) as a linear function of $EPUC_{t-1}$ and Q_{t-1} .

²¹ The logistic growth function is given by

$$\Delta X(t)/X(t-1) = v(1 - X(t-1)/K) + \epsilon(t);$$

where v and K are parameters. Subtracting the Schaefer production function gives a net change in CPUE of

$$CPUE(t)/CPUE(t-1) = (1 + v) - vCPUE(t-1)/K g_0 - g_0 E(t-1) + \epsilon(t);$$

where $E(\cdot)$ is aggregate effort. Ordinary least squares estimation then yields the estimate of g_0 .

Table 1. A Comparison of the Alternative Regimes

	First-Best	Comp.	Specific Tax	Quota
Effort	1,357.5	1,357.5	1,357.5	1,397.8
Fleet size	70.7	97.0	70.7	71.7
Tax rate			74.4	
Quota				482.5
Expected output	4,155.3	5,698.3	4,155.3	4,338.5
ENSB	5,819,967.3	5,787,239.2	5,819,967.3	5,815,285.5
CS	415,922.2	782,159.5	415,922.2	453,394.2
RENT	308,953.3	0	308,953.3	280,196.0
DEFNB	5,095,091.8	5,005,079.7	5,095,091.8	5,081,695.2

Note: *ENSB* corresponds to the presented discounted value of the fishery's expected net benefits to society in the various situations. *CS* measures the consumer surplus, and *RENT* the economic rent accruing during the initial period of the planning horizon. Finally, *DEFNB* refers to the present discounted value of the expected future net benefits to society, i.e.,

$$DEFNB = \rho V_{t+1}[X(t+1)],$$

so,

$$ENSB = CS + RENT + DEFNB.$$

or approximately 7.2% of the industry price. Interestingly, this magnitude is similar to the tax proposed by Pearce for the British Columbia salmon fishery. Given the unpopularity of the Pearce tax on salmon, it is likely that a 7.2% halibut tax would meet the same fate.

Of greater interest are the individual boats quotas. We estimate an optimal quota of 482.5 tons per boat, with per boat effort (1,397.8 skate-soaks) and the size of the fleet (71.7 boats) being marginally larger than first-best levels. The quota succeeds in removing 85.7% of the loss of expected net social benefits associated with competitive behavior in this fishery.²² The truncating nature of the quota implies that its imposition will lead to a reduction in effort per boat, *ceteris paribus*. Recognizing this, the regulatory authority will reduce the size of the fleet below the competitive level in order to increase the expected marginal value of effort, thereby inducing a secondary increase in effort. As the fishery moves from a competitive to a quota equilibrium, the increase in predicted effort is consequence of this secondary adjustment.²³

Given these results, the actual situation in the fishery is of obvious interest. While season controls were in effect prior to 1979–80, since then there has been limited-entry licensing. The fun-

damental weakness of this scheme is well known: it does not control the fishing capacity of the fleet. Pearce recognizes this problem and recommends that individual boat quotas be adopted. "The outstanding advantage of this approach is that it eliminates the basic cause of overcapacity in the fishing industry by removing the incentives of individual fishermen to protect and increase their share of the catch" (Pearce, p. 84). He also suggests that quota licenses be issued for a limited period. When that period has elapsed, the government would then hold an auction and sell the quota license to the highest bidder. In this way, the government would recoup the rent associated with the quota.

The small difference between taxes and quotas in these simulations is surprising, and we are led to question whether the structure of the model has forced this result. For instance, the assumption that β , θ , and ϵ are uncorrelated is quite strong: it is possible to envisage situations in which this would not follow. The inclusion of either inter- or intratemporal correlation, or both, is found to affect the differences in expected net social benefits when comparing the competitive, tax, and quota regimes.²⁴ For instance, increasing the intratemporal correlation coefficient between ϵ and θ from zero induces additional losses in expected net social benefits for both the competitive and quota cases. To be more specific,

²² Because the rents associated with the quota accrue to the fishermen, we would almost certainly have observed a decrease in the optimal season length if the model had been formulated to consider this issue.

²³ The second-best nature of the quota will, in general, prevent effort per boat from being at an optimal level.

²⁴ While it is feasible to adjust the analyses of the first-best, tax, and competitive cases to allow for nonzero correlation coefficients, we are only able to approximate their impact with quotas.

if ε and θ are perfectly correlated the loss in expected net social benefits arising from competitive behavior will increase to \$53,929.23 and a quota would then remove 85.2% of the loss. Of course, the possibility of intertemporal correlation also exists. Suppose that, in addition to the intratemporal correlation between ε and θ discussed above, ε follows a first-order autoregressive process, i.e., $\varepsilon(t) = \gamma\varepsilon(t-1) + \mu(t)$, where γ is set equal to 0.50, and $\mu(t)$ is a random variable. Then the loss is expected net social benefits arising from competitive behavior falls to \$46,277.67, and a quota removes 86.6% of the loss. Thus, the lack of both inter- and intratemporal correlation may have introduced bias into the results summarized in table 1, but the direction of this bias is unclear.

The particular forms chosen for the demand, cost, and growth functions also may have constrained the result. For instance, suppose that costs are a cubic rather than a quadratic function of effort. Sensitivity analysis reveals that, if marginal costs are an increasing (concave/convex) function of effort, then the losses associated with competitive behavior and the quota regime are (increased/decreased). The potential bias arising from the particular demand relationship which is adopted is equally unclear. Suppose, for instance, that b_0 and b_1 are adjusted so as to keep the price constant at a particular output level: the price elasticity of demand will then change, and we find that the losses associated with free entry and the quota are inversely related to the elasticity. An analogous exercise with respect to the growth function's parameters yields similar results. The original specification of the various functional relationships may thus have introduced bias into the analysis, but in an unknown manner.

Concluding Comments

This paper has contributed to the continuing research on regulatory methods for commercial fisheries under uncertainty. While the timing of decision-making guarantees that taxation will be first-best in a limited information sense in our model, the loss of expected net benefits associated with a quota are small. This finding is especially surprising because catch and kill are typically not identical in the presence of a quota as a result of the high mortality rate of fish returned to the ocean. Given the political difficulty that regulators face in imposing taxes on

catch, a properly designed quota may be the best practical alternative for fishery regulation.

[Received January 1990; final revision received July 1990.]

References

- Andersen, P. "Commercial Fisheries Under Price Uncertainty." *J. Environ. Econ. and Manage.* 9(1982):11-28.
- Andersen, P., and J. G. Sutinen. "Stochastic Bioeconomics: A Review of Basic Methods and Results." *Marine Resour. Econ.* 1(1984):117-36.
- Anderson, E. "Taxes vs. Quotas for Regulating Fisheries Under Uncertainty: A Hybrid Discrete-Time Continuous-Time Model." *Marine Resour. Econ.* 3(1986):183-207.
- Chow, G. C. *Analysis and Control of Dynamic Economic Systems*. New York: John Wiley & Sons, 1975.
- Clark, C. W. *Bioeconomic Modelling and Fisheries Management*. New York: John Wiley & Sons, 1985a.
- . "The Effect of Fishermen's Quotas on Expected Catch Rates." *Marine Resour. Econ.* 1(1985b):419-27.
- Cook, B. A. "Optimal Levels for Canada's Pacific Halibut Catch." Dep. Econ Disc. Pap. 1983-2, University of New Brunswick, June 1983.
- Cook, B. A., and P. Copes. "Optimal Levels for Canada's Pacific Halibut Catch." *Marine Resour. Econ.* 4(1987):45-61.
- Crutchfield, J. A., and A. Zellner. *Economic Aspects of the Pacific Halibut Fishery*. Fishery Industrial Research, vol. 1. Washington, DC: Bureau of Commercial Fisheries, 1962.
- Forsythe, G. E., M. A. Malcolm, and C. B. Moler. *Computer Methods for Mathematical Computations*. Englewood Cliffs NJ: Prentice-Hall, 1977.
- International Pacific Halibut Commission. Annual reports, Seattle.
- . *The Pacific Halibut: Biology, Fishery, and Management*. Seattle WA: IPHC Tech. Rep. No. 16, 1978.
- Koenig, E. F. "Controlling Stock Externalities in a Common Property Fishery Subject to Uncertainty." *J. Environ. Econ. and Manage.* 11(1984a):124-38.
- . "Fisheries Under Uncertainty: A Dynamic Analysis." *Marine Resour. Econ.* 1(1984b):193-208.
- . "Indirect Methods for Regulating Externalities Under Uncertainty." *Quart. J. Econ.* 99(1985):479-93.
- Myhre, R. *Minimum Size and Optimum Age of Entry for Pacific Halibut*. Seattle WA: IPHC Tech. Rep. No. 14, 1974.
- Pearse, P. H. *Turning the Tide: A New Policy for Canada's Pacific Fisheries*. Vancouver, British Columbia: Commission on Pacific Fisheries Policy, 1982.
- Schaefer, M. B. "Some Aspects of the Dynamics of Populations Important to the Management of Commercial Marine Fisheries." *Bull. Inter-Amer. Trop. Tuna Comm.* 1(1954):25-56.
- Stollery, K. R. "A Short-Run Model of Capital Stuffing in

- the Pacific Halibut Fishery." *Marine Resour. Econ.* 3(1986):137-53.
- Watson, L. T. "A Globally Convergent Algorithm for Computing Fixed Points of C^2 Maps." *Appl. Math. Computing* 5(1979):297-311.
- Watson, L. T., and D. Fenner. "Chow-Yorke Algorithm for Fixed Points or Zeros of C^2 Maps." *ACM Trans. on Math. Software* 6(1980):252-59.
- Weitzman, M. L. "Prices vs. Quantities." *Rev. Econ. Stud.* 41(1974):477-91.

Endangered Species and the Safe Minimum Standard

Richard C. Ready and Richard C. Bishop

The safe minimum standard of conservation is a decision rule for problems like endangered species preservation that involve irreversibility and uncertainty. It has been motivated as the minimax-loss solution to a two-person game against nature. However, two equally plausible games can be used to model decisions involving endangered species. For one game, the safe minimum standard, and therefore preservation, is preferred. For the other game, extinction is preferred. Although the safe minimum standard is intuitively appealing, it cannot be motivated by game theory.

Key words: conservation, endangered species, irreversibility, uncertainty.

Extinction of plants and animals is an example of a class of economic problems that involve irreversible destruction of environmental amenities and uncertainty over the future value of those amenities. If preserved, a species might become an important resource in the future. However, how valuable a particular species might become is at present unknown. At the same time, preservation of species may involve near-term costs. Ideally, economic analysis of issues involving endangered species would focus on the potential Pareto improvement criterion using cost-benefit analysis, including consideration of quasi-option value (Arrow and Fisher, Fisher and Hanemann). This approach may be stymied, however, by inability to assign probabilities to alternative outcomes or even to place bounds on potential payoffs from preservation (Smith and Krutilla, Bishop 1979).

Ciriacy-Wantrup proposed an alternative decision rule for problems involving irreversibility and pure uncertainty, the safe minimum standard of conservation (*SMS*). Wantrup was concerned with "flow resources with a critical zone," renewable resources where irreversible depletion is possible. Wantrup's concern was that the

irreversible loss of potential resources would limit society's ability to adapt to a changing world, leading to stagnation and the risk of "immoderate future losses." He argued that the appropriate policy for critical zone flow resources was to insure against such losses by adopting the safe minimum standard of conservation, which states that flows should be maintained at a level that makes it feasible to rebuild the stock in the future. For example, Ciriacy-Wantrup and Phillips stated that:

The objectives of conservation policy—and of many other public policies—can often be compared with the objectives of an insurance policy against serious losses that resist quantitative measurement. Here the objective is not to maximize a definite quantitative net gain but to choose premium payments and losses in such a way that maximum possible losses are minimized. (Ciriacy-Wantrup and Phillips, p. 28)

Wantrup personally applied the *SMS* concept to water quality (Ciriacy-Wantrup 1961) and agricultural land (Ciriacy-Wantrup 1964) as well as endangered species (Ciriacy-Wantrup 1952, Ciriacy-Wantrup and Phillips).

Whereas cost-benefit analysis has a sound theoretical foundation in the potential Pareto improvement criterion, the safe minimum standard has not been adequately tied to a theoretical model of social choice. Wantrup claimed that the safe minimum standard could be "regarded as a conceptual relative of the min-max solution or saddle point of a two-person strictly determined game" (Ciriacy-Wantrup 1968, p. 89), but he never formalized this relationship. Bishop (1978) attempted to formalize the relationship between

Richard C. Ready is an assistant professor, Department of Agricultural Economics, University of Kentucky. Richard C. Bishop is a professor, Department of Agricultural Economics, University of Wisconsin.

Journal Paper No. 90-1-73 of the Kentucky Agricultural Experiment Station. This research was supported by the Kentucky Agricultural Experiment Station and the College of Agricultural and Life Sciences, University of Wisconsin.

The authors thank Angelos Pagoulatos and the anonymous referees for helpful comments on earlier drafts.

the SMS and the minimax solution to a two-person game, but he incorrectly specified the loss matrix for the game he envisioned.¹

The purpose of this paper is to explore the link between game theory and the safe minimum standard of conservation. This will be done in the context of species preservation, but the results are applicable for a wide range of decisions involving irreversible destruction of natural resources. We identify two separate games that could be used to model endangered species preservation decisions, the insurance game and the lottery game. While either model is plausible, they yield completely different conclusions regarding endangered species policy.

The Insurance Game

A highly simplified version of the problem helps focus attention on central issues. Suppose that society is confronted with whether or not to develop resources where development will entail the extinction of a species. The development benefits will be symbolized by B_d , where B_d is assumed net of any known losses in use or non-use values associated with the extinct species.² Nevertheless, $B_d > 0$ is assumed to hold. Otherwise, extinction would never be efficient. Also, it will be assumed that preservation of the species requires only that development be abandoned; there are no out-of-pocket costs for preservation.

Following Wantrup, extinction is assumed to entail the possibility of future losses, symbolized by L , which are assumed to be large; L is at least large enough so that $L > B_d$. For example, L might be the losses incurred from the possible, but uncertain, outbreak of some deadly disease like AIDS. If the species in question is present, a cure for the disease will be found with certainty, making losses zero. However, if the species is extinct, the disease will, if it does occur, inflict heavy losses on society, measured by L . Uncertainty is modeled by two policy-independent future states of the world, state 1 and state 2, corresponding to an outbreak of the disease and no outbreak, respectively. We assume

that L and B_d are known, but that the probabilities of state 1 and state 2 are unknown.

In game-theoretic terms, society has two possible strategies, *SMS* and *DEV*, representing the maintenance of the safe minimum standard and development, respectively. The baseline for measurement of losses is arbitrarily chosen as a situation of no disease and no development. The loss matrix for the insurance game is constructed by measuring gains and losses from that baseline and is presented in table 1. Each row depicts a social strategy, and each column represents a state of the world. In the upper left-hand corner, the *SMS* is chosen and an outbreak of the disease occurs. Because the species is present, the disease is cured and no losses are sustained. Thus, the loss, relative to the baseline, is zero. If the *SMS* were chosen and the disease did not occur (the upper right-hand corner), the loss would again be zero. Choosing development implies a gain of B_d (a loss of $-B_d$). If development is chosen and an outbreak of the disease does not occur, then the total loss is simply $-B_d$. If development is chosen and the outbreak does occur, then the total loss is $L - B_d$, which is positive by definition. This is the immoderate future loss that Wantrup wanted to avoid. If this situation is viewed as a two-person, zero-sum game played against nature, the strategy that minimizes that maximum possible loss is preferred. Under the *SMS*, the maximum possible loss is 0. Under *DEV*, the maximum possible loss is $L - B_d$. The preferred strategy is then *SMS*.

Wantrup's idea of insurance against large future losses is implemented in that choosing the *SMS* insures that such losses will be avoided. Unfortunately, the conclusion that the *SMS* is the recommended policy for endangered species decisions is not robust to the specification of the game. To illustrate this, a different game, called the lottery game, is presented.

The Lottery Game

The assumptions are changed only slightly. Suppose that the same endangered species and the same disease are involved. In the insurance game, we assumed that the endangered species held the key for curing the disease, and that the outbreak of the disease was uncertain. In the lottery game, the disease is assumed to occur with certainty, but there is uncertainty about whether preservation of the species will lead to a cure. In state 1, the species does hold the cure, in state 2, it does not. An example is a disease such as

¹ In his loss matrix, Bishop (1978) double counted the benefits from species preservation in his state 2.

² Following Krutilla, non-use values are values that are held by individuals above and beyond values associated with personal use of the species. To the extent that such values are motivated by benevolence toward future generations (Krutilla's "bequest values"), they may incorporate consideration of the avoidance of the loss L , leading to some double counting in our game matrix. This issue will be ignored here in order to concentrate on the nature of the game.

Table 1. Loss Matrix for the Insurance Game

Strategy	State		Max Loss
	1	2	
<i>SMS</i>	0	0	0
<i>DEV</i>	$L - B_d$	$-B_d$	$L - B_d$

cancer, that we know will occur. The endangered species could be a plant that might provide chemical compounds helpful in treating cancer.

The baseline for the lottery game is no development and no cure. The loss matrix for the lottery game is presented in table 2. If the *SMS* is chosen and state 1 occurs, the losses from the disease are avoided, giving net losses to society, relative to the baseline, of $-L$ (a gain). If the *SMS* is chosen and state 2 occurs, the baseline situation occurs with net losses of 0. If *DEV* is chosen, the disease will not be cured regardless of the state because the species is unavailable. Net losses are $-B_d$ (a gain) in either state.

The lottery game captures an often-heard argument for species preservation: that endangered species should be saved because they may lead to the discovery of important new resources. Preserving the species is analogous to buying a lottery ticket. Adopt the *SMS* and society may (or may not) win big.

The maximum possible loss under the *SMS* in the lottery game is 0, while under *DEV* it is $-B_d$, which is less than zero. The minimax strategy for the lottery game therefore is to choose development, and guarantee a gain of B_d , even though it means extinction of the species and thus insures that society will suffer the losses associated with the disease. Thus, a seemingly innocuous modification of the motivation of the problem has completely reversed the conclusion.³

Table 2. Loss Matrix for the Lottery Game

Strategy	State		Max Loss
	1	2	
<i>SMS</i>	$-L$	0	0
<i>DEV</i>	$-B_d$	$-B_d$	$-B_d$

³ A similar reversal of the conclusion would occur in the insurance game if a third state of the world were introduced, where an outbreak of the disease occurs but the species does not provide the cure.

Taking Stock

Neither game captures the nature of the problem better than the other. Wantrup clearly viewed species preservation according to the insurance game. Bishop tended to motivate the problem along the lines of the lottery game but incorrectly specified the loss matrix, leading him to conclude the *SMS* was the preferred strategy. While Fisher and Hanemann operated in a cost-benefit framework, they discussed aspects of both games, and focused attention on a problem that is best modeled by the lottery game. Here, we have shown that, depending on how the endangered species problem is motivated, the conclusions reached using a minimax loss decision rule could be completely reversed. Use of a game-theoretic framework gives structure to the problem. However, *SMS* cannot be motivated as the minimax-loss solution to a two-person game against nature.

Interestingly, using the minimax-regret decision rule, the conclusion is that the *SMS* should be adopted for either the lottery game or the insurance game. There is some evidence that individuals consider regret when making decisions under uncertainty, rather than maximizing expected utility (Loomes and Sugden). Perhaps the intuitive belief that endangered species should be preserved is motivated by regret considerations. However, while individuals may make decisions based on regret, it is questionable whether public decisions regarding endangered species should be made that way.

Rejection of game theory does not necessarily imply rejection of the *SMS* as a decision-making policy, although it does leave us without a theoretical foundation for that policy. Society may choose to adopt the *SMS* not because it results from a rigorous model of social choice, but simply because individuals in the society feel that the *SMS* is "the right thing to do." Indeed, the Endangered Species Act of 1966 (with subsequent amendments) seems to codify that feeling, allowing only a limited role for economic analysis in endangered species decision making.

[Received September 1989; final revision received May 1990.]

References

- Arrow, K. J., and A. C. Fisher. "Environmental Preservation, Uncertainty, and Irreversibility." *Quart. J. Econ.* 88(1974):313-19.
- Bishop, R. C. "Endangered Species and Uncertainty: The

- Economics of a Safe Minimum Standard." *Amer. J. Agr. Econ.* 57(1978):10-18.
- . "Endangered Species, Irreversibilities, and Uncertainty: A Reply." *Amer. J. Agr. Econ.* 58(1979):376-79.
- Ciriacy-Wantrup, S. V. *Resource Conservation: Economics and Policies*. Berkeley: University of California Press, 1952; and 3rd ed. 1968.
- . "The 'New' Competition for Land and Some Implications for Public Policy." *Nat. Resour. J.* (1964):252-67.
- . "Water Quality, A Problem for the Economist." *J. Farm Econ.* 43(1961):1133-44.
- Ciriacy-Wantrup, S. V., and W. E. Phillips. "Conservation of the California Tule Elk: A Socioeconomic Study of a Survival Problem." *Biolog. Conserv.* 3(1970):23-32.
- Fisher, A. C., and W. M. Hanemann. "Option Value and the Extinction of Species." *Advances in Applied Micro-Economics*, vol. 4., ed. V. K. Smith. Greenwich CT: JAI Press, 1986.
- Krutilla, J. V. "Conservation Reconsidered." *Amer. Econ. Rev.* 57(1967):777-86.
- Loomes, G., and R. Sugden. "Regret Theory: An Alternative Theory of Rational Choice Under Uncertainty." *Econ. J.* 92(1982):805-24.
- Smith, V. K., and J. V. Krutilla. "Endangered Species, Irreversibilities, and Uncertainty: A Comment." *Amer. J. Agr. Econ.* 58(1979):371-75.

Demands for Local Public Sector Outputs in Rural and Urban Municipalities

Melville L. McMillan and Joe Amoako-Tuffour

Little is known about local public services demands in rural municipalities and how they differ from demands in urban municipalities. Local public sector output demands in rural and urban municipalities in Victoria, Australia, are studied here using a system of equations approach based on the translog model. Expenditure, price, and substitution elasticities are reported. Demands differ between rural and urban municipalities and even among the (especially urban) subclasses.

Key words: Australia, local government, municipal service demands, rural, urban.

Many studies have considered the demands for municipal government outputs, but these studies have focused almost exclusively on urban municipalities (e.g., Bergstrom and Goodman; Deacon; Ladd; Hayes 1985, 1986; Hayes and Grosskopf; Hayes and Slottje; McMillan, Wilson, and Arthur; Perkins). With rare exceptions (Kiefer, McMillan 1985), rural municipalities have been ignored. Indeed, it is unusual that differences between types of municipalities are central to any such analysis (Hayes 1985; McMillan, Wilson, and Arthur). The objective of this paper is to compare the demands for local public services in rural jurisdictions with those in urban municipalities. The data encompass all municipalities in the State of Victoria, Australia. Municipalities there range from densely populated cities in the Melbourne metropolitan area to sparsely populated rural municipalities.

Demands for local public services may differ between rural and urban municipalities for various reasons. The spatially dispersed nature of rural households and business is expected to make

road transportation a priority in order to service rural enterprises, market their products, and provide access to services available only in urban centers. A sparse population also reduces the potential economies of public services like water, sewerage, and garbage and results in more private provision. The higher travel costs incurred to use centrally located public recreational and cultural facilities may reduce the demand for them. The importance of property ownership in rural municipalities and local governments' reliance on property taxes may imply a perceived greater property tax burden in rural communities and a more conservative attitude toward local government spending. Also, the relative numbers and the political weight of households in rural and urban portions of non-homogenous municipalities may influence demand. The data studied here indicate distinctly different expenditure patterns between rural and urban local governments. The objectives are to determine whether such differences warrant disaggregation of municipalities by type when examining local public service demands and, if so, what differences in demand exist.

The characteristics of the demands for six municipal services are reported for rural and urban municipalities. While attention is focused on these two groups, the rural group is actually the combination of two types of rural municipalities (rural shires and rural with small urban centers) and the urban group is a composite of three types of urban municipalities (small independent urban, suburban, and central metro-

Melville L. McMillan is a professor, Department of Economics, University of Alberta; and Joe Amoako-Tuffour is an assistant professor, Department of Economics, St. Francis Xavier University, Antigonish, Nova Scotia.

The authors thank Adolf Buse, David Ryan, and Alan Sharpe for their valuable comments and assistance. They also thank the Central Research Fund of the University of Alberta for its support. For their support in the collection and initial assessment of the data, McMillan thanks the Social Science and Humanities Research Council of Canada, the Centre for Research on Federal Financial Relations at the Australian National University, and the Victoria Grants Commission. The authors also thank *Journal* referees for their constructive comments.

politan cities). Thus, we also provide the results for each subgroup and briefly consider within rural and within urban subgroup differences.¹ The nature of preferences for municipal services based on all 211 Victoria municipalities together is reported in McMillan and Amoako-Tuffour (1988b).

In the following section, the characteristics of the municipal classes are discussed. The indirect translog model used to estimate the demand functions then is outlined and the data are discussed. Next, the empirical results and the estimated elasticities are reported by municipality type. A conclusion completes the paper.

Municipality Classification and Characteristics

The mean values of the local characteristics for the various classes are reported in table 1. When considering the two group rural-urban division, in general, the rural municipalities have much smaller populations, lower per capita incomes, a larger tax base per capita but a smaller portion being residential property, greater local expenditures per capita (notably for roads), and a less dense and less urbanized population than the urban municipalities. Comparing the two rural categories, the rural shires have smaller populations and less (indeed almost no) population growth, a smaller share of their population is in urban areas (10% as opposed to 37%), and, although both types of rural municipalities typically incorporate several population centers, the average population of each center is 65 in the rural shires in contrast to 330 in the other group.

Three subgroups constitute the urban municipalities. One group is essentially the cities in the central portion of the Melbourne area plus two nearby cities; it is designated as the central metropolitan cities group. These cities are large, densely populated, house a somewhat older population, and are experiencing population decline. Provincial cities dominate the small, independent urban center group. The municipalities in this group are typically separated from other urban centers and, relative to other urban

municipalities, have smaller populations, lower household incomes, and higher per capita local government expenditures. Predominately suburban municipalities, many from the Melbourne metropolitan area, form the remaining urban group. This group has large and relatively young populations, multiple population centers, rapid growth, and high household incomes.

Public expenditure characteristics in these groups exhibit distinctly different patterns. Per capita expenditures in rural municipalities average \$247 in contrast to \$165 in urban communities. Budget allocations are also quite different. Expenditures for general government and road maintenance represent almost two-thirds of the rural budget but less than one-third of urban. Road maintenance represents the greatest difference—less than one-tenth of urban outlays but one-third in rural municipalities generally and four-tenths in the rural shire subgroup. Together, expenditures for education, health and welfare, recreation and culture, and community amenities and economic services are one-fourth of rural budgets but one-half of urban outlays. Although smaller proportionately, the per capita expenditures for education, health, and welfare differ little between rural and urban municipalities. The same is true of "other" expenditures. The budget shares devoted to expenditure areas vary substantially between rural and urban municipalities and, to a lesser extent, among their subgroups. The analysis below seeks to explain that variation as a function of the relative cost of services in the communities and the expenditure level.

Because Victoria's municipal governments are not responsible for schooling, police protection, or social welfare, their role and budgets are more modest than is usually the case in the United States or Canada. Otherwise, the municipal authorities provide a conventional range of municipal services. Municipalities are governed by locally elected councils. In 1978–79, property ownership or occupancy restrictions implied that not everyone could vote in local elections. There is now universal franchise with special provisions for nonresident property owners. (For further discussion of the data and the institutional environment, see McMillan 1985.)

The Model

The indirect utility function is the foundation of this analysis. Utility is a function of commodity prices, P_i and the expenditure, M , on the bundle

¹ Cluster analysis was used to divide the 211 Victoria municipalities into the five classes studied here (McMillan 1983). Shire is the Victoria term for a rural municipality and is used here to designate the most rural subset of that group. In 1978, there were 133 shires in Victoria, most of which fall into the 111 rural group with almost all the others, on the fringe of major urban areas, fitting into the suburban category.

Table 1. Characteristics of Categories of Victoria Municipalities (mean values)

	Rural Shires (51 mun.)	Rural with Decentralized Urban Centers (60 mun.)	Rural (Combined) (111 mun.)	Small Independent Urban (35 mun.)	Suburban (38 mun.)	Central Metropolitan Cities (27 mun.)	Urban (Combined) (100 mun.)
Population	3,229	5,746	4,589	11,941	43,658	46,840	33,416
Population change, 1974-79(%)	0.04	7.9	4.3	5.8	16.1	-8.1	5.9
Urban population (%)	10.0	37.3	24.7	96.4	86.4	100.0	93.6
Density (pop/km ²)	0.026	0.064	0.045	5.5	5.5	30.3	12.2
Number of population centers	5.0	6.5	5.7	1.3	3.1	1.0	1.9
Value of property per capita	1,437	1,054	1,230	953	946	1,151	1,004
Residential share of value	0.115	0.315	0.223	0.710	0.682	0.711	0.699
Commercial-industrial share	0.033	0.074	0.055	0.237	0.189	0.289	0.233
Rural share of value	0.852	0.611	0.722	0.052	0.219	0.001	0.068
Expenditure per capita (\$)	282.78	217.24	247.4	192.46	144.87	157.58	164.9
Share of expenditure on:							
GEN ^a	0.311	0.303	0.307	0.233	0.243	0.214	0.231
ROAD	0.406	0.283	0.340	0.096	0.093	0.058	0.085
EHW	0.051	0.087	0.070	0.106	0.105	0.132	0.113
REC	0.082	0.136	0.111	0.213	0.247	0.243	0.234
CAES	0.054	0.075	0.065	0.168	0.153	0.215	0.175
OTH	0.097	0.115	0.107	0.184	0.157	0.138	0.162
Non-tax revenue per capita (\$)	90.08	61.21	74.47	40.29	27.13	27.84	311.93
Income per household	7,109	7,744	7,468	8,130	11,329	10,150	9,091
Income less than \$5,000 (%)	36.5	31.8	33.9	31.4	15.3	24.5	23.4
Income more than \$15,000 (%)	6.7	7.8	7.3	8.1	17.2	14.5	13.3
Workforce without skills (%)	76.1	70.9	73.3	67.9	65.6	63.7	65.9
Professions (%)	6.8	8.8	7.9	10.8	11.5	15.1	12.2
Age less than 20 (%)	39.0	35.4	38.7	35.4	41.0	27.6	35.4
Age more than 65 (%)	9.6	10.1	9.8	13.0	5.4	12.5	9.8

* The designations refer to expenditure categories as follows: (GEN) general government, (ROAD) road maintenance, (EHW) education, health and welfare, (REC) recreation and culture, (CAES) community amenities and economic services, and (OTH) other public services. Expenditure shares may not sum to 1.0 because of rounding errors.

of goods. The translog flexible functional form advanced by Christensen, Jorgenson, and Lau was chosen for estimation.

$$(1) \quad \ln V = \alpha_0 + \sum_i \alpha_i \ln(P_i/M) + \frac{1}{2} \sum_i \sum_j \beta_{ij} \ln(P_i/M) \ln(P_j/M).$$

The translog function serves as a local approximation of the exact (twice differentiable) indirect utility function about a point. Using Roy's Identity and maximizing utility subject to the budget constraint $M = \sum P_i X_i$, yields demand equations in expenditure share form:

$$(2) \quad \frac{P_i X_i}{M} = \frac{\alpha_i + \sum_j \beta_{ij} \ln(P_j/M)}{\sum_j \alpha_j + \sum_j \sum_i \beta_{ij} \ln(P_j/M)} \quad (i = 1, 2, \dots, n)$$

when $\beta_{ij} = \beta_{ji}$ in the indirect utility function (1) above.

For consistency with the assumptions of utility maximization subject to a budget constraint, (a) the substitution matrix must be symmetric (requiring $\beta_{ij} = \beta_{ji}$) and negative semidefinite, and (b) the demand functions must be homogeneous of degree zero in prices and expenditure which requires that $\sum \alpha_i = 1$ (a normalization) and $\sum_j \beta_{ij} = 0$ for each i . We maintain the symmetry restrictions and the normalization on α but test for homotheticity.

The system of equations to be estimated is limited to the set of local publicly provided commodities. In doing so, the demands for local public commodities are assumed a weakly separable branch of the utility function. This assumption is consistent with a two-stage decision-making process in which local households collectively decide how much to allocate to the local public sector and then how to allocate it among the various local public outputs. That collective provision is utility maximizing is maintained by the usual assumption (Bergstrom and Goodman) that public output is consistent with the preferences of the median voter.

One can speculate as to whether a utility maximization model is appropriate for the study of public outputs. Strong conditions are necessary if even aggregate private good data are to satisfy the predictions of consumer choice theory. The political processes involved in translating individual voter's preferences into collective choices are more complex and presumably argue for an

even more tenuous association with utility maximization. On the other hand, utility maximization based on the median voter model is now central to the theory of local public provision and has seen successful empirical applications (see Inman and Rubinfeld). Deacon, for example, provides a strong argument in favor of the utility-maximizing model. However, when the above conditions necessary for constrained utility maximization are tested, they are rarely fully satisfied. But then, similar tests of private good demands are frequently unsuccessful. On a more positive note, using a nonparametric test based on the generalized axiom of revealed preference, DeBoer and, to a somewhat lesser degree, McMillan and Amoako-Tuffour (1988b) provide evidence suggesting that the constrained maximization model may characterize the allocation of public expenditures.

Although estimation of flexible functional forms imposes minimal *a priori* restrictions on the underlying preferences, few studies have used this method in estimating the demands for public outputs (e.g., Dunne and Smith; Dunne, Pashardes and Smith; Hayes and Grosskopf; Grosskopf and Hayes 1983, 1986; McMillan and Amoako-Tuffour 1988b). Also, the single-equation approach usually used (Deacon and Perkins are exceptions) has meant that potentially important complementary and substitute relationships have generally been neglected. The system of equations resulting from the translog model employed here permits these to be recognized fully—and, as a result of the rural-urban disaggregation, even interjurisdictional differences. The indirect specification allows us to use information available on the prices of local public services in Victoria's municipalities.

Data

Systems of demand equations were estimated using 1978–79 cross-section data for the 211 municipal governments in Victoria, Australia. Measures of the costs (or prices) of local public services are a unique feature of the data. These measures are derived from unpublished data of the Victoria Grants Commission. As part of its procedure, the commission estimates the expenditure required for each municipality to provide a standard level of service. Allowance is made for intercommunity variations in (a) the number of units to be served ("expenditure needs") and (b) the local unit cost of the service ("expenditure disability"). Expenditure needs arise from

deviations from the standard measure of units to be served. For example, a disproportionately high share of population under the age of five implies greater expenditures for health services, and more kilometers of roads per property impose extra road costs. A disability allowance adjusts for unusual unit costs due to special difficulties in providing local services. The estimated expenditure requirements are discounted to reflect only the expenditure financed from the municipality's own revenues.² In 1980, assessments were made for eleven types of expenditures, which in the collection and analysis of the data were selectively combined into six expenditure categories: general administration (27.8% of the total); road maintenance (24.7%); education, health, and welfare (8.3%); recreation and culture (15.6%); community amenities and economic services (largely garbage disposal and street cleaning and lighting) (10.5%); and other (13.1%). The price or cost for each service is defined as the expenditure (specifically the amount financed locally) deemed by the commission as necessary to provide the standard level of service divided by the municipal population. This figure does not represent the unit cost of the service, but rather, the average price to the local taxpayer. Thus, the per capita cost of services as derived from the commission's estimates are relatively greater where the commission found exceptional expenditure needs or expenditure disabilities.

Such a measure of the relative price of public services is quite unusual, especially for local governments. (Deacon, and Blackley and DeBoer also established relative prices.) Typically in demand for local public good studies, the price to the decisive voter is based upon grants, tax share, and/or possible tax exporting.³ Assuming other governments maintain their sharing ratios, grants are already reflected in the price indexes. In this analysis, local tax share is uniform for all six categories of public services, so it does not affect relative prices. Similarly, although McMillan (1985) found evidence in this data of perceived tax shifting/exporting, such evidence does not change relative prices within a municipality and therefore will not influence budget shares. Thus, for an analysis of demand based on expenditure shares, relative price measures (as outlined here) are both necessary and ade-

quate in the context of the usual public sector prices.

Demographic variables (e.g., community characteristics) could also be incorporated in the model, e.g., Dunne and Smith. The work of Hayes (1986) and McMillan, Wilson, and Arthur suggests adding population. The extent to which population and other demographic characteristics reflect expenditure needs has already been incorporated into the price term. However, demographic variables might be added to pick up interjurisdictional variations in the private cost of consuming public services or to recognize possible taste differences. Efforts to incorporate demographic variables into this model encountered difficulties. However, McMillan (1985) found demographic variables in this data not very important in explaining per capita expenditures. Furthermore, Barnes and Gillingham found the introduction of demographic variables in a system of consumer demand equations inferior to using unpooled data. The division of this data into rural and urban (and the associated subgroups) achieves a degree of unpooling by demographic characteristics expected to be related to costs and tastes and rationalizes not including demographic variables here. Not incorporating demographic variables into the model results in the specification employed here being consistent with our earlier work.

General Empirical Issues

The demand system was estimated for each of the seven municipal classifications in table 1. Each system has six local public good categories, but restrictions require that only five equations be estimated.⁴ The equation for the other expenditure category (*OTH*) was excluded leaving the five equations for general government (*GEN*), road maintenance (*ROAD*), education, health, and welfare (*EHW*), recreation and culture (*REC*), and community amenities and economic services (*CAES*).

The system was first estimated with symmetry (i.e., $B_{ij} = B_{ji}$, and also the normalization $\sum \alpha_{ij} = 1$ and the equality constraint) imposed and again with the homogeneity constraint $\sum_j B_{ij} = 0$ added. The estimated parameters are not re-

² In determining its grant recommendations, the commission also considers revenue-raising capacity, population, and area.

³ See Inman and Rubinfeld for discussion. Alternatively, some use wage rates and labor shares (e.g., Hayes).

⁴ Because the budget shares must sum to unity, the disturbances are not independent so the six-equation variance-covariance matrix is singular. This problem is handled by arbitrarily omitting one equation. The system of the $n - 1$ remaining equations was then estimated using SHAZAM (version 5.0).

ported here but are available in McMillan and Amoako-Tuffour (1988a). The log likelihood values and test statistics of homogeneity conditional on symmetry are reported in table 2. The test rejects the homogeneity restrictions, a condition required for utility maximization, for all municipality classes but not for the small independent urban group. Although homogeneity is not rejected for one class, for consistency with the other estimates, the following results are from the models without the homogeneity restrictions.

The log likelihood values are also used to test the disaggregation of the municipalities into the various subgroups. Those tests indicate that the proposed subdivisions are justifiable (table 2). Informational inaccuracy values (Theil and Mnookin) support the disaggregation. The informational inaccuracy value (the mean accuracy) typically diminishes with disaggregation. Although public service price variations explain a large part of budget share differences among the total group, the greater homogeneity within subgroups reduces the overall inaccuracies of the estimates (McMillan and Amoako-Tuffour 1988a). These results substantiate the earlier conjectures that demands for municipal services are not alike among municipalities of different types. While the rural-urban separation was expected, finding the distinctions among the subgroups was surprising.

The expenditure and the uncompensated own-price and cross-price elasticities for the seven municipality groupings are reported in tables 3, 4, and 5. Elasticities were evaluated at the sample means using the elasticity formulas found in Christensen and Manser. The Allen partial elas-

ticities of substitution are reported in table 6, and the Morishima elasticities of substitution for the combined rural and combined urban groups are reported in table 7.

A Rural-Urban Comparison

The empirical results are discussed first for the two class rural-urban division. The discussion then turns to the disaggregation of the rural municipalities and finally for the subdivision of the urban municipalities.

Expenditure Elasticities

Except for the relatively low expenditure elasticity for other expenditures (about 0.3) for both the rural and urban groups, the expenditure elasticities in those two groups in table 3 range from about 0.6 to 1.6. The recreation and culture category has the largest expenditure elasticity at approximately 1.6 for both the rural and urban groups. Although there is some difference between the expenditure elasticities in the two groups, the levels are quite similar except for the community amenities category which has a value of about 1.27 for the rural group but 0.75 for the urban group. The nature of other expenditures, which include debt servicing and repair and restoration after natural disasters, likely explains its relatively low elasticity. This low value, as well as the range in the other expenditure elasticities reflects the merits of not imposing homogeneity which would restrict these elasticities to unity.

Table 2. Tests of the Homogeneity Restriction (log-likelihood values)

	Rural Shires	Rural with Decentralized Urban Centers	Rural (Combined)	Small Independent Urban	Suburban	Central Metropolitan Cities	Urban (Combined)
Symmetry ^{a,d}	504.9	577.6	1039.1	326.1	428.4	308.5	960.4
Symmetry and homogeneity	495.0	560.2	1020.9	321.6	396.6	296.0	940.2
Test statistic ^{b,c} (-2LLF)	19.8	34.8	36.4	10.0	63.6	25.0	40.4

^a The normalization $\sum \alpha_i = 1$ and the equality constraint ($\sum \beta_{ij}$ is the same in each equation) are also maintained along with the symmetry restriction, $\beta_{ij} = \beta_{ji}$.

^b The test statistic is minus twice the log of the likelihood function; i.e., $-2LLF = -2(\ln L_R - \ln L_U)$, where L_R and L_U are the likelihood values of the restricted (homogeneity and symmetry) and the unrestricted (symmetry only) models.

^c Six additional restrictions are required for homogeneity. The critical χ^2 values at the 5% and 1% level of significance with six degrees of freedom are 12.59 and 16.81, respectively.

^d These values are used to test the various disaggregations. For the division of all municipalities into the rural group and the urban group, for example, yields the test statistic 197.6 for 20 degrees of freedom from $-2\ln \lambda = [1039.1 + 960.4] - 1900.7$. The χ^2 test for further disaggregation of the combined rural and combined urban cases yields 85.2 and 205.2, respectively. The critical χ^2 values at the 5% level are 31.41 and 55.76 for 20 and 40 degrees of freedom.

Table 3. Expenditure Elasticities of Demand

Municipality Class	Budget Category ^a					
	GEN	ROAD	EHW	REC	CAES	OTH
Rural shires	0.832	1.098	0.490	1.830	1.448	0.483
Rural with decentralized urban	1.013	1.000	1.012	1.821	0.572	0.271
Rural (combined)	0.884	1.146	0.616	1.594	1.269	0.340
Small independent urban	1.385	0.369	1.111	1.445	0.662	0.571
Suburban	0.720	0.837	0.439	2.231	0.901	0.060
Central metropolitan cities	1.180	0.919	1.343	1.396	0.555	0.419
Urban (combined)	1.076	1.113	0.903	1.602	0.745	0.303

^a Elasticities are calculated at class mean values.

Table 4. Uncompensated Own-Price Elasticities of Demand

Municipality Class	Budget Category ^a					
	GEN	ROAD	EHW	REC	CAES	OTH
Rural shires	-0.623 ^{*b}	-0.879*	-0.848*	-0.925*	-0.639*	0.296
Rural with decentralized urban	-0.440*	-0.756*	-1.103*	-1.247*	-0.413*	-0.108
Rural (combined)	-0.567*	-0.719*	-0.881*	-0.932*	-0.781*	0.036
Small independent urban	-0.384	-0.414*	0.506	-0.272	-0.349*	-0.221
Suburban	-1.237*	-0.231*	-0.621	-0.880	0.284	-0.046
Central metropolitan cities	-0.630	-0.733*	-1.165	-0.387	-0.206	-0.398
Urban (combined)	-0.851*	-0.444*	0.046	-0.438	-0.291*	-0.129

^a Elasticities are calculated at class means.

^b Asterisk indicates an elasticity at least twice the standard error of the elasticity of the model with symmetry and homogeneity imposed

Deacon's study is the only other one familiar to us which reports expenditure elasticities for local public outputs. (Blackley and DeBoer, and Dunne and Smith report budget/expenditure elasticities for various state-local and national expenditures.) Although the functions he examines differ from the categories used in this study, with one exception (municipal courts at 2.10) the expenditure elasticities vary from 0.85 to 1.08. Income elasticities are commonly reported and are typically less than one (Inman). Expenditure elasticities of about one are not inconsistent with income elasticities of less than one because of the intervening effect of the elasticity of total local government expenditures with respect to income, a value which is typically less than one (Inman).

Price Elasticities

Uncompensated own-price elasticities of demand are reported in table 4. Unlike the expenditure elasticities, the own-price elasticities for the two municipal groups are less similar, appearing more inelastic for the urban municipalities with the exception of the general govern-

ment category. One positive own-price elasticity occurs in each group, but the values are close to zero. Only elasticities with the appropriate negative sign are believed likely to differ significantly from zero.⁵ For the two aggregated rural and urban classes, no own-price elasticity exceeds (absolute) 1.0.

Price elasticities reported in single-equation studies of local public services are generally inelastic, but they vary substantially across functions and among studies, perhaps partly because of differences in the definition of the price term. Comparison with other studies is impeded because the responsibilities of Victoria's local governments do not include schooling and po-

⁵ Elasticities in tables 4, 5, and 6 which are designated with * are at least twice the standard error of the elasticity estimated from the model with both symmetry and homogeneity imposed. The estimated elasticities are quite similar for the two models. With homogeneity, however, calculation of the standard errors of the estimated elasticities is straightforward (Binswanger) and is believed to be a reasonably reliable indicator of significance in this instance. Evidence on confidence intervals for the elasticities of substitution support this position (McMillan and Amoako-Tuffour 1988a). The significance of the expenditure elasticities cannot be approximated this way because they are constrained to 1.0 by the homogeneity assumption.

Table 5. Uncompensated Cross-Price Elasticities of Demand

Municipality Class and Budget Category	Budget Category ^a					
	GEN	ROAD	EHW	REC	CAES	OTH
Rural shires						
GEN		-0.064	-0.017	-0.090	-0.076* ^b	-0.080
ROAD	-0.037		-0.021	0.057	-0.095*	-0.123
EHW	0.001	0.081		-0.115	0.391*	0.000
REC	-0.657*	-0.014	-0.140		0.372*	-0.465
CAES	-0.633*	-0.859*	0.322*	0.597*		-0.236
OTH	-0.151	-0.267	0.000	-0.281	-0.079	
Rural with decentralized urban						
GEN		-0.116*	-0.182*	-0.124	-0.065	-0.084
ROAD	-0.119		0.020	0.115	-0.138*	-0.121
EHW	-0.631*	0.063		0.585	0.266	-0.192
REC	-0.520	0.007	0.306		-0.127	-0.238
CAES	-0.128	-0.401	0.346*	-0.060		0.084
OTH	0.003	-0.092	-0.081	-0.070	0.077	
Rural (combined)						
GEN		-0.030	-0.069	-0.096	-0.059	-0.064
ROAD	-0.107*		-0.064*	-0.031	-0.092*	-0.132
EHW	-0.218	-0.129		0.424	0.259*	-0.070
REC	-0.482*	-0.247*	0.201		0.165	-0.300
CAES	-0.394*	-0.523*	0.233*	0.316*		-0.120
OTH	-0.016	-0.148	-0.027	-0.171	-0.013	
Small independent urban						
GEN		-0.070	-0.416	-0.336	-0.069	-0.110
ROAD	0.066		0.098	-0.080	-0.000	-0.039
EHW	-0.847	0.017		-0.418	-0.039	-0.331
REC	-0.382	-0.139	-0.244		-0.235	-0.172
CAES	0.073	-0.028	0.023	-0.130		-0.250
OTH	0.051	-0.040	-0.134	-0.134	-0.213	
Suburban						
GEN		0.137*	-0.263	0.733	-0.014	-0.075
ROAD	0.327*		-0.034	-0.583*	-0.132	-0.182
EHW	-0.542	0.007		0.116	0.399	0.201
REC	0.350	-0.350*	-0.139		-0.820*	-0.393
CAES	-0.066	-0.087	0.226	-1.002*		0.255
OTH	0.045	-0.035	0.174	-0.078	-0.118	
Central metropolitan cities						
GEN		0.008	0.101	-0.471	-0.025	-0.163
ROAD	0.085		0.109	-0.106	-0.303*	0.028
EHW	0.129	0.023		0.095	-0.173	-0.253
REC	-0.461	-0.053	0.045		-0.501*	-0.039
CAES	0.109	-0.061	-0.002	-0.363*		-0.032
OTH	-0.091	0.041	-0.121	0.169	-0.020	
Urban (combined)						
GEN		0.065	-0.188	0.053	-0.073	-0.082
ROAD	0.168*		-0.096	-0.380*	-0.188	-0.173
EHW	-0.346	-0.055		-0.460	0.125	-0.214
REC	-0.069	-0.179	-0.300*		-0.381	-0.234
CAES	-0.020	-0.060	0.099	-0.310*		-0.162
OTH	0.062	-0.022	-0.082	-0.034	-0.098	

^a Elasticities are calculated at class means.^b Asterisk indicates an elasticity at least twice the standard error of the elasticity of the model with symmetry and homogeneity imposed.

lice and fire protection. Even if one considers estimates of the elasticity of demand for parks and recreation (which should parallel the recreation and culture category here), the price elasticities of demand reported in Inman range from -0.19 to -0.92. In his earlier investigation

of this data, McMillan (1985) estimated expenditure responses to price variables implying price elasticities of demand of -0.1 for total municipal expenditures, -0.2 for property-related expenditures and -0.4 for personal services. Because of the smaller potential for interfunction substi-

Table 6. Allen Partial Elasticities of Substitution

Municipality Class and Budget Category	Budget Category ^a					
	GEN	ROAD	EHW	REC	CAES	OTH
Rural Shires						
GEN	-1.183* ^b	0.979*	0.494	-0.283	-0.589	-0.004
ROAD	0.979*	-1.067*	0.690	1.795*	-0.666	-0.176
EHW	0.494	0.690	-16.222*	-0.926	7.788*	0.490
REC	-0.283	1.795*	-0.926	-9.518*	8.771*	-2.969
CAES	-0.589	-0.666	7.788*	8.771*	-10.476*	-0.984
OTH	-0.004	-0.176	0.490	-2.969	-0.984	3.534
Rural with decentralized urban						
GEN	-0.446	0.605*	-1.079	0.096	0.149	0.279
ROAD	0.605*	-1.654*	1.232*	1.845*	-0.837	-0.051
EHW	-1.079	1.232*	-11.628*	5.326	4.538*	-0.654
REC	0.096	1.845*	5.326	-7.377*	0.130	-0.249
CAES	0.149	-0.837	4.538*	0.130	-4.904*	1.299
OTH	0.279	-0.051	-0.654	-0.249	1.299	-0.667
Rural (combined)						
GEN	-0.966*	0.797*	-0.096	0.019	-0.018	0.287
ROAD	0.797*	-0.967*	0.235	0.869*	-0.265	-0.096
EHW	-0.096*	0.235	-11.881*	4.449*	4.569*	-0.039
REC	0.019	0.869*	4.449*	-6.830*	4.121*	-1.208
CAES	-0.018	-0.265	4.569*	4.120*	-10.671*	0.144
OTH	0.287	-0.096	-0.039	-1.208	0.144	0.675
Small independent urban						
GEN	-0.263	0.653	-2.527	-0.194	0.973	0.788
ROAD	0.653	-3.954*	1.295	-0.009	0.368	0.156
EHW	-2.527	1.295	5.872	-0.853	0.879	-0.686
REC	-0.194	-0.009	-0.853	0.165	0.050	0.508
CAES	0.973	0.368	0.879	0.050	-1.412*	-0.698
OTH	0.788	0.156	-0.686	0.508	-0.698	-0.631
Suburban						
GEN	-4.369*	2.181*	-1.790	3.673*	0.628	0.243
ROAD	2.181*	-1.638*	0.513	-1.514*	-0.029	-0.319
EHW	-1.790	0.513	-5.486	0.908	3.054	1.716
REC	3.673*	-1.514*	0.908	-1.317	-3.139*	-0.257
CAES	0.628	-0.029	3.054	-3.139*	2.760	-0.714
OTH	0.243	-0.319	1.716	-0.257	-0.714	-0.232
Central metropolitan cities						
GEN	-1.759	1.316	1.945	-0.756	1.062*	-0.002
ROAD	1.316	-11.739*	1.748	0.484	-0.492	1.125
EHW	1.945	1.748	-7.473	1.736	0.538	-0.494
REC	-0.756	0.484	1.736	-0.196	-0.939	1.115
CAES	1.062*	-0.492	0.538	-0.939	-0.405	0.324
OTH	-0.003	1.125	-0.494	1.115	0.324	-2.468
Urban (combined)						
GEN	-2.598*	1.840*	-0.592	1.303	0.658	0.569
ROAD	1.840*	-4.118*	0.258	-0.509	0.036	0.045
EHW	-0.592	0.258	1.310	-1.057	1.620*	-0.421
REC	1.303	-0.509	-1.057	-0.270	-0.579	0.156
CAES	0.658	0.036	1.620*	-0.579	-0.921*	-0.257
OTH	0.569	0.045	-0.421	0.156	-0.258	-0.493

^a Elasticities calculated at the class mean values.^b Asterisk indicates an elasticity at least twice the standard error of the elasticity of the model with symmetry and homogeneity imposed.

tution, it is reasonable that the elasticities of aggregated bundles of services would generally be (absolutely) smaller than those for the individual services.

Cross-price elasticities (table 5) are typically low. For these two groups, only one exceeds

0.5 (absolute) and almost one-half are less than 0.1 (absolute). This result is consistent with those of the two other studies reporting cross-elasticities, Ehrenberg and Perkins, but contrast with those of Dunne, Pashardes, and Smith, for United Kingdom government expenditures.

Table 7. Morishima Elasticities of Substitution

Municipal Class and Budget Category	Budget Category					
	GEN	ROAD	EHW	REC	CAES	OTH
Rural (combined)						
GEN		0.603	0.815	0.748	0.733	0.043
ROAD	0.537		0.838	0.843	0.713	0.010
EHW	0.264	0.412		1.253	0.998	0.015
REC	0.294	0.621	1.143		0.991	-0.109
CAES	0.350	0.307	1.097	1.170		0.019
OTH	0.397	0.350	0.828	0.638	0.728	
Urban (combined)						
GEN		0.506	-0.215	0.368	0.276	0.172
ROAD	1.028		-0.119	-0.056	0.167	0.087
EHW	0.465	0.371		-0.184	0.444	0.012
REC	0.903	0.306	-0.267		0.060	0.105
CAES	0.754	0.353	0.035	-0.072		0.038
OTH	0.733	0.353	-0.195	0.100	0.116	

Elasticities of Substitution

The Allen partial elasticities of substitution are reported in table 6.⁶ For the urban group of municipalities, general government expenditure and expenditures in the various other expenditure categories (except for education, health and welfare) are substitutes. An especially high degree of substitution is observed with road maintenance. The substitutability between general government and road maintenance also exists, although at a more modest level, in the rural group, and is the only notable interdependency between general government and the other five expenditure areas. For the rural group, the large positive elasticities (all over 4.0) indicate a high degree of substitutability among education, health and welfare, recreation and culture, and community amenities and economic services.⁷ Within the urban group, outlays on education, health and welfare, and on community amenities and economic services are also substitutable ($\sigma_{EHW, CAES}^A = 1.62$), but the other two relationships are complementary. Otherwise, the elasticities of substitution are generally low for the two municipal groups suggesting that strong substitution or complementary relationships do not exist between local public sector outputs.

The Allen partial elasticity of substitution has

⁶ Each matrix is symmetric about the diagonal. For both groups, one value of σ_{ij}^A is greater than zero, indicating that convexity conditions are not satisfied at the point of estimation. Also, the monotonicity condition is not satisfied for these rural data.

⁷ These elasticities are surprisingly large. Deacon and Perkins observed substitution effects between fire and police protection. Perkins also finds some interdependency between health and sanitation outlays and between parks and recreation and libraries; these results conform with those observed here.

been criticized (e.g., Blackorby and Russell) as an imperfect measure of the elasticity of substitution. The crux of the problem is that the Allen elasticity, which is the ratio of the compensated cross-price elasticity to the cost share (i.e., $\sigma_{ij}^A = \eta_{ij}^c/s_j$), measures only how one commodity adjusts to the change in another's price. When there are more than two inputs, this does not allow for the adjustment of other commodities. Blackorby and Russell advocate the Morishima elasticity as more economically meaningful. The Morishima elasticity of substitution, which equals the difference between the compensated cross and own-price elasticities (i.e., $\sigma_{ij}^M = \eta_{ij}^c - \eta_{jj}^c$), measures how the ratio of commodities adjusts to a change in the price of commodity j . (See Chambers, chap. 3 for further discussion and illustration.) In order that the reader may better understand substitutability among commodities and appreciate the difference between the two measures, we report the Morishima elasticities for the rural and urban classifications in table 7. A complete set is available from the authors.

Morishima elasticities are generally not symmetric (i.e., the value depends upon which price changes) and Morishima substitutability/complementarity need not match Allen substitutability/complementarity. For the rural municipalities, only one elasticity ($\sigma_{REC, OTH}^M$) indicates complementarity, while ten did according to the Allen elasticities. In the case of $\sigma_{GEN, j}^M$, general government now substitutes for all other budget categories, but the elasticities of substitution for *EHW*, *REC*, and *CAES* are now somewhat larger. When the price of general government changes, $\sigma_{i, GEN}^M$, the relationship is somewhat stronger for

road maintenance (as before). With Morishima elasticities of 0.99 or larger, the substitutability among health, education and welfare, recreation and culture, and community amenities and economic services is less than reflected by the Allen measure but is still pronounced.

For the urban group, the extent of complementarity is again reduced by the Morishima measure but, as before, is concentrated in the education, health and welfare, and the recreation categories. Now, however, with the lack of symmetry, complementarity is concentrated only in the columns (i.e., $\sigma_{i, EHW}^M$ and $\sigma_{i, REC}^M$), where all but one of the signs match those of the Allen elasticity. General government and road maintenance continue to show greater degrees of substitutability with each other. Although the Morishima elasticities are smaller, the relationships among the education, health and welfare, the recreation and culture, and the community amenities and economic services categories are retained. Overall, while the Morishima elasticities tend to be smaller and indicate less complementarity than the Allen measures, important parallels remain in the two sets of results.

Two Types of Rural Municipalities

The rural group examined above is the combination of a group of rural municipalities with limited urban populations (rural shires) and rural municipalities with more significant urban development. For the two subgroups, the expenditure elasticities parallel each other except for the *EHW* and the *CAES* categories. In the case of education, health and welfare, the expenditure elasticity for rural shires, at about 0.5, is half that for the rural decentralized urban group; but for community amenities, at 1.4, it is about twice as large. The own-price elasticities are consistent for the two groups except that the value for the other expenditure category for the rural shires has a positive sign. The price interdependence between community amenities and economic services and various other expenditure categories for the combined rural municipalities seems to stem from a strong relationship in the rural shires. There, eight of ten of those cross-price elasticities are likely significantly different from zero, and six of those eight have a value in excess of (absolute) 0.5.

The Allen elasticities of substitution reveal some interesting differences. For all expenditure categories but road maintenance, the elasticities with general expenditures differ in sign between

the rural shires and the rural municipalities with larger urban populations, i.e., when they are substitutes in one class they are complements in the other. The substantial three-way substitutability found among the *EHW*, *REC*, and *CAES* categories in the combined rural group now separates into two patterns. In the more urbanized rural municipalities, a high degree of substitutability still exists between education, health and welfare and recreation and culture and community amenities; however, substitutability is not maintained between recreation and culture and community amenities. For the rural shires, the high elasticities of substitution between community amenities and education, health and welfare, and between community amenities and recreation and culture remain, but that between education, health and welfare and recreation and culture does not. Also, unlike in the aggregate rural case, road maintenance and the two personal services categories (*EHW* and *REC*) exhibit a considerably greater degree of substitutability in these two classes.

The model for the rural shires does not meet convexity conditions ($\sigma_{OTH, OTH}^A > 0$), but the model for the more urbanized rural municipalities does satisfy convexity. Also, monotonicity is not satisfied for the rural shires.

The Urban Municipal Groups

Three types of urban municipalities were identified in the five class categorization of the Victoria municipalities—central metropolitan cities, small independent urban municipalities, and suburban municipalities. With the exception of road maintenance, the expenditure elasticities for the central metropolitan cities and the small independent urban groups closely parallel each other.⁸ The expenditure elasticities for the suburban group tend to differ from the other two—those for community amenities and (especially) for recreation and culture are higher while those for the other expenditure category, for education, health and welfare, and (to a lesser extent) for general government are lower.

Only for the central metropolitan cities group are all own-price elasticities negative. Positive price elasticities occur in the suburban group for community amenities and in the small city group for the education, health, and welfare category

⁸ Recall that the homogeneity restriction, implying expenditure elasticities of 1.0, could not be rejected for the small independent urban group.

but are not believed to differ significantly from zero. Demands are generally price inelastic. In only two cases are values in excess of (absolute) 1.0 recorded, and only seven of eighteen exceed (absolute) 0.5. However, these elasticities vary considerably across municipal groups. Other than those for road maintenance, unlike the results for the rural categories, few elasticities are likely to differ significantly from zero.

The cross-price elasticities are generally low for these municipalities too. In only one case, between education, health and welfare, and recreation and culture in suburban municipalities, does the value reach one ($\eta_{CAES, REC} = 1.002$). Even when the values are relatively large, there is little consistency or pattern among the different municipal groups.

For the three classes of urban municipalities, a high degree of substitutability or complementarity (using the Allen measure) exists among expenditure categories, especially for the suburban and central metropolitan groups for which almost one-half the elasticities of substitution exceed absolute 1.0. However, except for the suburban group, few of these elasticities are likely to differ significantly from zero. The interdependence is particularly noticeable among general government and the other five budget areas. Only for general government and road maintenance and for general government and community amenities is the substitute relationship consistent across all three municipal groups. These elasticities are also quite large over the three municipal groups for the education, health, and welfare group; but, again, the signs are not uniform. Interdependencies occur among education, health and welfare, recreation and culture, and community amenities, but they are not consistent across municipalities.

Convexity conditions are not satisfied for any of the urban municipal groups, but monotonicity is met for all three groups.

Conclusions

Very little is known about the demands for local public services in municipalities of different types, and particularly about those in rural municipalities. This is so although there are various reasons for expecting demands to differ, especially between rural and urban jurisdictions. A system of translog equations was used to estimate the demands for six local public sector outputs in the kinds of municipalities spanning the full range found in Victoria, Australia. It was

found that demands differ between urban and rural municipalities and, somewhat surprisingly, even among the five categories comprising these two groups, at least to the extent that municipal budget shares can be explained by the local costs of provision and total expenditure.

The demand differences which exist are revealed in the elasticities and are usually most obvious between the rural and urban groups. For example, own-price elasticities tend to be more inelastic for urban municipalities and interdependencies between expenditure categories are stronger in the rural than in the urban (except for the suburban) groups.

Demand elasticities are both intellectually interesting and policy relevant. They are helpful to those attempting to explain or predict municipal expenditure changes. Price and expenditure elasticities are especially valuable to those projecting the impacts of tax changes, grants designed to promote certain kinds of expenditures, or grants to equalize fiscal capacities. Measures recognizing interdependencies among budget categories, as reported here, are also useful in efforts to predict the full implications of fiscal reforms such as the "New Federalism" (Craig and Inman), determine cost of public service indices (Deacon), or define more general measures of fiscal capacity (Akin).

The elasticities reported here are generally consistent with the (admittedly quite broad range of) elasticities reported in studies of local (and typically urban) government outputs. At the same time, the demands do vary among municipalities of different types. This result suggests that some of the observed variations in elasticities may result from differences in the type of municipalities, even among urban municipalities, studied. It also implies that greater attention should be paid to distinguishing the types of municipalities surveyed and that more analysis could be done of public sector demands in the classes of municipalities now underrepresented in the literature.

[Received August 1989; final revision received August 1990]

References

- Akin, John S. "Estimates of State Resource Constraints Derived from a Specific Utility Function: An Alternative Measure of Fiscal Capacity." *Nat. Tax J.* 32(1979):61-71.
- Barnes, Roberta, and Robert Gillingham. "Demographic

- Effects in Demand Analysis: Estimation of the Quadratic Expenditure System Using Microdata." *Rev. Econ. and Statist.* 64(1984):591-601.
- Bergstrom, T. C., and R. P. Goodman. "Private Demands for Public Goods." *Amer. Econ. Rev.* 63(1973):280-97.
- Binswanger, Hans P. "A Cost Function Approach to the Measurement of Factor Demand and Elasticities of Substitution." *Amer. J. Agr. Econ.* 56(1974):377-86.
- Blackley, Paul R., and Larry DeBoer. "Measuring Basic Wants for State and Local Public Goods: A Preference Independence Transformation Approach." *Rev. Econ. and Statist.* 64(1987):418-25.
- Blackorby, Charles, and R. Robert Russell. "Will the Real Elasticity of Substitution Please Stand Up? *Amer. Econ. Rev.* 79(1989):882-88.
- Chambers, Robert G. *Applied Production Analysis: A Dual Approach*. London: Cambridge University Press, 1988.
- Christensen, L. R., D. W. Jorgenson, and L. J. Lau. "Transcendental Logarithmic Utility Functions." *Amer. Econ. Rev.* 65(1975):367-83.
- Christensen, L. R., and M. E. Manser. "Cost-of-Living Indexes for U.S. Meat and Produce, 1947-1971." *Household Production and Consumption*, ed. N. E. Terleckyj. New York: National Bureau of Economic Research, 1975.
- Craig, Steven G., and Robert P. Inman. "Education, Welfare and the 'New' Federalism: State Budgeting in a Federalist Public Economy." *Studies in State and Local Public Finance*, ed. H. S. Rosen. Chicago: University of Chicago Press, 1986.
- Deacon, R. T. "A Demand Model for the Local Public Sector." *Rev. Econ. and Statist.* 60(1978):184-92.
- DeBoer, Larry. "State and Local Government Utility Maximization According to GARP." *Public Finan. Quart.* 14(1986):87-99.
- Dunne, J. P., P. Pashardes, and R. P. Smith. "Needs, Costs and Bureaucracy: The Allocation of Public Consumption in the U.K." *Econ. J.* 94(1984):1-15.
- Dunne, J. P., and R. P. Smith. "The Allocative Efficiency of Government Expenditure." *Eur. Econ. Rev.* 20(1983):381-94.
- Ehrenberg, R. G. "The Demand for State and Local Government Employees." *Amer. Econ. Rev.* 63(1973):366-79.
- Grosskopf, S., and K. Hayes. "Do Local Governments Maximize Anything?" *Public Finan. Quart.* 11(1983):202-16.
- . "The Demand for Local Public Goods: Choosing the Appropriate Functional Form." *Appl. Econ.* 18(1986):1179-92.
- Hayes, Kathy. "Congestion Measures for Local Public Goods in Metropolitan and Nonmetropolitan Cities." *Growth and Change* 16(1985):1-9.
- . "Local Public Good Demands and Demographic Effects." *Appl. Econ.* 18(1986):1039-45.
- Hayes, Kathy, and Shawna Grosskopf. "The Role of Functional Form in Estimating the Demand for Public Goods." *Rev. Econ. and Statist.* 66(1984):169-73.
- Hayes, Kathy, and Daniel Slottje. "Measures of Publicness Based on Demographic Scaling." *Rev. Econ. and Statist.* 69(1987):713-18.
- Inman, Robert P. "The Fiscal Performance of Local Governments: An Interpretative Review." *Current Issues in Urban Economics*, ed. P. Mieszkowski and M. Straszheim. Baltimore, MD: Johns Hopkins University Press, Baltimore, 1979.
- Kiefer, David. "The Dynamic Behavior of Public Budgets: An Empirical Study of Australian Local Governments." *Rev. Econ. and Statist.* 63(1981):422-29.
- Ladd, H. F. "Municipal Expenditures and the Composition of the Local Property Tax Base." *Property Taxation, Land Use and Public Policy*, ed. A. D. Lynn, Jr. Madison: University of Wisconsin Press, 1976.
- McMillan, Melville L. "Tax Base Composition, Tax Prices and Fiscal Equalization Policy: the Case of Victoria, Australia." *Perspectives on Local Public Finance*, ed. J. M. Quigley. London: JAI Press, 1985.
- . *The Categorization of Local Authorities: An Experiment with Victoria Municipalities*. Dep. Econ. Res. Pap. No. 83-7, University of Alberta, April 1983.
- McMillan, Melville L. and Joe Amoako-Tuffour. *An Examination of Demands for Local Public Sector Outputs in Different Types of Municipalities*. Dep. Econ. Res. Pap. No. 88-10, University of Alberta, 1988a.
- . "An Examination of Preferences for Local Public Sector Outputs." *Rev. Econ. and Statist.* 70(1986):45-54.
- McMillan, Melville L., W. Robert Wilson, and Louise M. Arthur. "The Publicness of Local Public Goods: Evidence from Ontario Municipalities." *Can. J. Econ.* 14(1981):596-608.
- Perkins, George M. "The Demand for Local Public Goods: Elasticities of Demand for Own Price, Cross Prices and Income." *Nat. Tax J.* 30(1977):411-22.
- Rubinfeld, Daniel L. "The Economics of the Local Public Sector." *Handbook of Public Economics*, vol. 2, ed. Alan J. Auerbach and Martin Feldstein. Amsterdam: North-Holland Publishing Co., 1987.
- Theil, H., and Robert H. Mnookin. "The Informational Value of Demand Equations and Predictions." *J. Polit. Econ.* 74(1966):34-45.

Distributional Effects of Household Linkages

G. Andrew Bernat, Jr. and Thomas G. Johnson

The study uses input-output models with disaggregated household sectors to demonstrate that important interhousehold linkages are ignored in standard input-output models of both rural and urban regions. The results indicate that higher-income households are not strongly linked to low-income households and that, given equal changes in exogenous income, low-income households will have much higher impacts on the local economy than high-income households. In addition, significant differences in household linkages were found between rural and urban regions.

Key words: household income, linkages, regional input-output models.

Knowledge of the linkages among economic sectors has long been employed in regional impact models, most notably in input-output (I-O) models. The level of sectoral disaggregation varies widely from model to model, with many models including considerable detail for at least some of the goods and service producing sectors. However, despite recent advances in modeling the household sector, most regional I-O models still treat the household sector as a single, aggregated sector. This treatment masks the complex linkages both among household groups within the household sector as well as between these household groups and other economic sectors.

The purpose of this paper is to demonstrate the large differences in household linkages in small, rural regions. Large differences in linkages according to household income indicate the need for models that can capture these differences (Johnson and Capps; Martin and Henry; Stevens and Trainer; Park, Mohtadi, and Kurbursi). The distribution of household linkages determines how a change in the income of one household group affects the distribution of income among all households. For example, trickle-down theory, in which spending by high-in-

come households leads to higher incomes for lower-income households, requires strong linkages between high-income households and low-income households. Finally, differences in linkages according to the level of development of an economy are important to consider as well. In order to analyze these issues, household linkages in four regions, three rural and one urban, of Virginia are examined.

The concept of household linkages is developed in the next section. The models are then discussed and the results presented, followed by a summary and conclusion.

Household Linkages

Two economic sectors are directly linked if one sector purchases goods or services from the other. Total linkages consist of both direct linkages and indirect linkages generated by the intersectoral relationships in an economy.

Households are linked to one another through income generation and induced spending. Two major characteristics of interhousehold linkages are magnitude (or strength) and distribution. Magnitude is the income received by the second household from an exogenous increase in the income of the first household. The total strength of linkages for a household is the sum of all increases in household income resulting from the initial increase.

The strength of linkage is greater for some households than others. This pattern of varying strength of household linkages reflects the dis-

G. Andrew Bernat, Jr., is an agricultural economist with the Economic Research Service, U.S. Department of Agriculture; Thomas G. Johnson is an associate professor, Virginia Polytechnic Institute and State University.

The views expressed in this paper are those of the authors and do not necessarily reflect the policies of the U.S. Department of Agriculture.

The authors gratefully acknowledge helpful comments from Ken Hanson, Chinkook Lee, and Gerald Schluter.

tribution of linkages. In a uniform distribution, the strength of linkage from a particular household is the same to each other household; thus, a change in the income of one household would result in an equal change in the income of all other households. In contrast, nonuniform linkages would result in larger changes for some households than others. In an extreme form, trickle-down theory asserts that the linkages from upper- to lower-income households are stronger than the linkages from upper-income households to other upper-income households. Clearly, the distribution of linkages among households indicates how a change in the income of a particular household will affect the overall distribution of income.

If the linkages among households are not uniform, a change in the income of one household would have a different effect on the distribution of income than an equal change in the income of another household. Household linkages are not necessarily symmetric. Household *X* may be strongly linked to household *Y*, while household *Y* is weakly linked to household *X*.

The strength of total linkages between any two household groups is a function of three types of linkage: (a) the linkage between households and producing sectors as embodied in the household consumption function (consumption linkages); (b) the linkages among the various producing sectors of the region (interindustry linkages); and (c) the linkage between producing sectors and households (value-added linkages). The dependence of indirect household linkages on these three basic linkages implies that we should observe certain patterns in the strength and distribution of household linkages from region to region, as explained below.

First, the marginal propensity to consume (MPC) is usually considered to be a declining function of household income. Thus, a given marginal change in household income will result in a smaller total change in expenditures for high-income households than for low-income households. In a regional context, particularly for rural regions, the decline in MPC with income will be magnified if the regional share of the household expenditure bundle also declines as household income increases.

Second, if high-income households spend relatively more on services than do lower income households, and if the production of services is low-wage-intensive, high-income households would have stronger linkages to low-income households than would low-or middle-income households. A recent study of household link-

ages in West Virginia (Rose and Beaumont 1988) showed no support for trickle-down theory. Because of the similarities between West Virginia and the regions studied here, we also anticipate weaker linkages from upper-income households to lower-income households than from upper income households to other households.

Third, household linkages will be relatively weak in economies with weak interindustry linkages. With few interindustry linkages, a relatively small proportion of any change in the output of one sector will be transmitted to other local sectors. Thus, even if consumption linkages are relatively strong, a large portion of any increase in household expenditures will flow out of the region as demand for intermediate goods and services. Because rural regions have relatively weak interindustry linkages compared to urban regions of similar size, stronger household linkages are expected in more urban versus rural regions.

Finally, the dependence of household linkages on both value-added linkages and interindustry linkages means that linkages are distributed more uniformly, the stronger are interindustry linkages. If interindustry linkages are relatively weak, the distribution of household linkages for a given household group will be dominated by the distribution of value-added of the producing sectors to which the households have strong consumption linkages. If differences in value-added linkages across sectors are large, and if households have different patterns of consumption linkages, the distribution of household linkages will differ markedly across household groups. On the other hand, in an economy with strong interindustry linkages, a few sectors will no longer account for most of the value-added generated. As intersectoral linkages increase, the distributional impacts of each household group will converge to an overall weighted average of the value-added linkages of all sectors in the region.¹ If more developed, urban economies do have stronger and more extensive interindustry linkages than less developed, rural economies, household linkages will be more uniform in more urban economies than in rural economies.

To summarize, the expected pattern of household linkages within and among regions is expressed by four hypotheses.² First, total house-

¹ This is similar to Weisskopf and Wolff's concept of the convergence of inequality impacts.

² These hypotheses indicate the anticipated results of this study. However, they cannot be tested statistically because of the type of input-output models employed in the study.

hold linkages decrease with household income. Second, the distribution of household linkages differs among household groups. Specifically, the linkages of upper-income households are skewed toward middle- and upper-income households. Third and fourth, household linkages in rural economies are relatively weak and less equally distributed than the linkages in urban regions.

Modeling Household Linkages

Measurement and comparison of household linkages requires explicit modeling of the expenditure patterns and income flows of different household groups. The standard input-output model is inadequate because it treats all households identically. Therefore, the household sector of the standard Type II regional input-output model must be expanded. Following Miyazawa (see also Johnson and Capps; Martin and Henry; Rose and Beaumont 1989, 1988; Rose, Stevens, and Davis; Schultz; Weisskoff; Weisskoff and Wolff), such a model can be represented by

$$(1) \quad \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} A & C \\ V & O \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} f \\ g \end{bmatrix},$$

where x is the vector of output levels of the n producing sectors in the region, A is the standard $n \times n$ matrix of technical coefficients, and f is the vector of final demand (excluding household demand) for the n producing sectors. This model departs from the standard I-O models by grouping the household sector into k income groups. Total household income, y , and exogenous households income, g , are represented by $k \times 1$ vectors rather than scalars. The consumption coefficients are represented by the $n \times k$ matrix C and the household value-added (income) coefficients are represented by the $k \times n$ matrix V .³

Solving (1) for x and y gives the matrix of multipliers:

$$(2) \quad \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} I - A & -C \\ -V & I \end{bmatrix}^{-1} \begin{bmatrix} f \\ g \end{bmatrix}.$$

Again following Miyazawa, (2) can be expressed in partitioned form:

$$(3) \quad \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} B(I + CKVB) & BCK \\ KVB & K \end{bmatrix} \begin{bmatrix} f \\ g \end{bmatrix},$$

where $B = (I - A)^{-1}$ and $K = (I - VBC)^{-1}$. Matrix B is the multiplier matrix of a Type I input-output model and indicates the change in the output of each producing sector required for a unit change in final demand of each producing sector, excluding the demand generated by households. The expression in the upper left portion of the multiplier matrix indicates the total change in sectoral output resulting from a change in final demand, including the induced effects of households. The elements of BCK , the upper right-hand portion, indicate the change in regional output that is attributable to an exogenous change in the income of each of the k household groups. The elements of KVB measure the change in the income of each of the k household groups due to a change in final demand (vector f) of each good and service producing sector.

Finally, the $k \times k$ matrix K indicates the direct plus indirect changes in income for each household group that would result from an exogenous change in income for each household group. The matrix K thus measures the multiplier effects (or linkages) within the household sector and is therefore the focus of our study. Each column of the matrix K shows the multipliers for a different household group. The first row shows the effects on the lowest income group, while the last row shows the effects on the highest income group.

The size and pattern of the elements of K indicate the linkage among households. Differences in the elements in any column j of the matrix K indicate a nonuniform distribution of linkages from households in group j to other households. For example, larger elements for low-income rows than for middle- or upper-income rows in the high-income columns of the matrix K would be consistent with trickle-down theory.

If the columns are not identical, linkages differ across household groups. Varying strengths of linkage across household groups are indicated if the elements in any row differ across columns. If the column totals differ, then the strength of total linkages differs by income. The sum of these totals is the analog of the household income multiplier of the standard Type II input-output model (Miyazawa).

Model Implementation

The study regions were chosen to compare different types of rural regions and to compare rural and urban regions. In particular, the regions

³ This formulation assumes no direct transfers between households. To model such transfers, the lower right-hand diagonal submatrix would be nonzero.

were chosen to represent a wide spectrum of industrial compositions and of income distributions. These considerations led to the selection of four regions in Virginia.

Regions 1, 2, and 4 are rural whereas region 3 is highly urban. The coal industry is the dominant industry in region 2, while regions 1 and 4 are primarily agricultural. Differences in either the strength or distribution of household linkages that result from differences in sectoral composition should be apparent because of the differences in economic structure represented by these regions. In addition, the hypothesis that the distribution of household linkages is a function of the degree of interindustry linkage, and thus regional development, can be examined by comparing regions 2 (the least developed region) and 3 (the most developed region). The distribution of income in region 4 is one of the most unequal in Virginia, whereas region 3 has one of the most equal distributions. Inequality in regions 1 and 2 is near the statewide average. These regions therefore permit examination of the relationship between inequality and the pattern of household linkages.

Each of the four models required the construction of matrices A , C , and V .⁴ The technical coefficients matrices, A , were constructed from a seventy-two-sector 1977 U.S. I-O model using the regional purchase coefficient method proposed by Stevens et al., of the Regional Science Research Institute (RSRI). The matrices of consumption coefficients, C , were based upon the Consumer Expenditure Surveys (U.S. Department of Agriculture, Bureau of Labor Statistics 1977a, b). The expenditure data for each of 12 household groups were converted from the 147 major commodity groups to the I-O sectors using a bridge matrix that was part of an input-output model of Virginia constructed by RSRI. The V matrix is based upon the Public Use Microdata (PUMS) tape from the 1980 Census of Population and Housing, a matrix of employment coefficients from the 1977 RSRI input-output model of Virginia, and Virginia Employment Commission data.

Analysis and Results

The most important characteristic of the K matrices to examine is whether the strength of household linkages differs across income groups.

Total household linkage for each household group is the column sum of the household multiplier matrix K . In all four regions, these totals are much lower for the higher-income groups than for the lower-income groups (table 1).

The differences in total household linkage presented in table 1 imply that aggregation of households into a single sector likely will introduce serious errors in the resulting multipliers, particularly when the impacts of exogenous changes in household income are considered. These results support the hypothesis that lower-income households are more strongly linked to the regional economy than are higher-income households.⁵

To facilitate comparisons of the distribution of household linkages, each of the four K matrices can be considered as a multiplier surface. The row and column axes represent household income groups, while the size of the elements of K represent the height of the surface. Each of the K matrices for the four regions exhibit certain regularities that can be quantified by estimating the following polynomial:

$$(4) \quad k_{ij} = b_0 + b_1i + b_2j + b_3i^2 + b_4j^2 + b_5ij + b_6i^3 + b_7j^3.$$

The dependent variable, the ij th element of K , is the change in income of the i th household group resulting from a dollar's change in the income of the j th household group. The row index of K is represented by i and the column index by j . The identity matrix was subtracted from each K matrix to allow examination of indirect effects alone. The diagonal elements of each multiplier matrix are composed of two parts. One part, called the direct effect, represents the tautology that an increase in the income of a household by a dollar raises the income of that household by a dollar. With no intersectoral or household linkages, the diagonal elements of these matrices would be unity. Because the immediate effect does not represent a linkage in any meaningful sense, it is removed by subtracting 1.0 from each element on the diagonal. The remaining portion of the diagonal elements measures the linkage of each household to itself. The results of estimating these equations using OLS are presented in table 2.

Equation (4) makes it possible to construct the multiplier surfaces and to clearly see how the distribution of the k_{ij} 's varies across rows, down

⁴ Further detail on both data sources and model construction is available on request and in Bernat (1985). Also see Bernat (1989) for an application of the model to income transfers.

⁵ This result clearly is a short-term result and ignores any long-run savings-investment linkages between households and the rest of the region.

Table 1. Total Linkages by Household Group

Group	Income (\$1,000)	Region 1	Region 2	Region 3	Region 4
1	<6	183.7	138.0	203.0	189.5
2	6-9	126.9	95.7	144.6	132.4
3	9-11	110.5	81.5	126.5	116.1
4	11-13	105.5	78.8	123.0	110.5
5	13-15	108.0	78.6	124.8	113.1
6	15-17	100.8	73.1	116.7	105.0
7	17-21	92.9	66.9	110.1	97.9
8	21-25	89.4	63.3	104.9	93.9
9	25-32	82.7	58.4	99.4	97.6
10	32-42	77.8	53.6	93.9	82.4
11	42-53	71.5	47.9	86.7	75.7
12	>53	49.8	33.0	60.6	52.0

Table 2. Household Multiplier Surface Equations

Variable	Coefficient	Region				Combined
		1	2	3	4	
Intercept	b_0	15.07** ^a	11.12**	23.22**	18.14**	16.89**
i	b_1	-3.37**	-2.88**	-14.02**	-3.66**	-5.98**
	b_2	-4.04**	-2.91**	-3.79	-4.17**	-3.73**
i^2	b_3	1.16**	0.90**	3.32**	0.99**	1.59**
j^2	b_4	-0.59**	0.44**	0.64	0.61**	0.52**
ij	b_5	-0.033*	-0.044**	-0.1*	-0.029**	-0.555**
i^3	b_6	-0.074**	-0.054**	-0.18*	-0.058**	-0.092**
j^3	b_7	-0.028**	-0.020**	-0.031	-0.029**	-0.027**
R^2		0.80	0.80	0.52	0.61	0.46

^a Single asterisk indicates significant at the 10% level; double asterisk indicates significant at the 1% level.

columns, and between regions. For example, the hypothesis that the distribution of linkages across households differs by income would be supported if the coefficients on the terms which include the column index (the j 's) are significantly different from zero. A uniform distribution of linkages among households would be indicated by insignificant coefficients on the terms which include the row index. Finally, a hill- or trough-shaped pattern is possible when the coefficients on either the row or column indices are of different signs. A positive b_3 and negative b_1 and b_6 would indicate that the elements of the columns of K rise and then decline as household income increases, depending on the relative magnitudes of the coefficients.

The results of estimating equation (4) for each region indicate that the K matrix for each region exhibits a strong hill-shaped pattern down the columns. The coefficients on i and i^3 are negative, while the coefficients on i^2 are positive, indicating that the elements of each column of K first increase and then decrease as one moves

from low-income rows to higher-income rows. Thus, household linkages were strongest to the middle- and upper-middle income groups for all households. Not only is there little evidence of strong trickle-down effects, but the hill-shaped pattern for every income group in each region indicates that an increase in household income of any household group would have a larger impact on middle- and upper-middle households than on either upper- or lower-income households.

A second feature of the K matrices is that the elements of each row decline as one moves from low-income columns to high-income columns. In other words, the strength of linkages to each household group declines with income. The positive coefficient on j^2 for some regions indicates an increase in multipliers at intermediate income levels, with declining multipliers at the highest income levels. These results are consistent with the hypothesis that high-income households are not strongly linked to low-income households.

As an aid in visualizing the patterns implied by the above equations, the multiplier surface for the combined regions is presented in figure 1. The vertical axis gives the value of the ij th element of matrix K . On the j axis is plotted the income group index of the spending households (columns of K) and on the i axis is the group index of the receiving households (rows of K). The rear left-hand corner of the diagram (the origin in three-dimensional space) represents both the row and column for the lowest household group. The height of the surface above any line parallel to axis j gives the size of the multipliers from households in group j to households in group i .

Support for the hypothesis of little or no trickle-down effect is indicated by the hill shape of the surface along the i axis in figure 1, a pattern shared by all regions. For each household group, linkages are weakest to the household groups 2, 3, and 4. As household income increases (the row index increases, represented by moving from left to right in figure 1), linkages strengthen,

peaking with household groups 8–10, and declining steeply for the highest two income groups. Thus, the strongest linkages for all households, and in all four regions, are with middle- and upper-middle-income-level households.

This finding is consistent with the Rose and Beaumont's results for West Virginia. Their results indicate that the strongest linkages for each household group are with households in the \$25,000 to \$35,000 and \$35,000 to \$50,000 income classes. We found the strongest links with households in the three classes between \$21,000 and \$42,000.

The results also support the hypothesis that the strength of household linkages is positively related to the level of economic development. For each household group, the highest total multipliers were for region 3, the most developed, urban region (table 1). Region 2 had the weakest household linkages; this region is dominated by coal mining and in many respects is the most rural and least developed of the four study regions. Linkages for low income house-

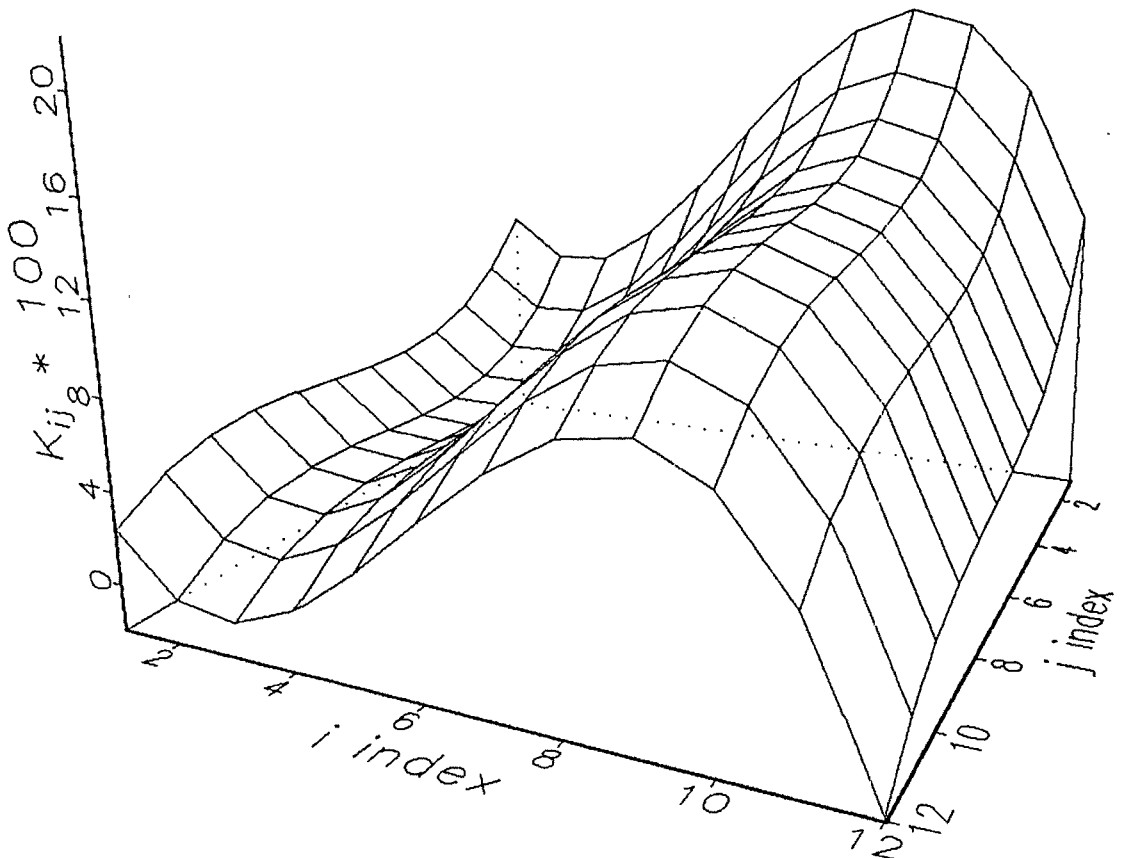


Figure 1. Multiplier surface for combined regions

holds in the coal region were 68% as strong as for region 3 and almost half for the highest income households. Total linkages in the two agriculture-based regions were similar to each other, ranging from about 92% of the region 3 linkages for low-income households to about 83% for high-income households.

In order to examine the hypothesis that household linkages are more uniform in more developed, urban regions, dummy variables for the urban region were added to the equation for the combined regions (d_1 , an intercept dummy; d_2 , d_3 , and d_4 are dummy interaction variables on i , i^2 , and i^3 ; and d_5 , d_6 , and d_7 are dummy interaction variables on j , j^2 , and j^3).⁶ The results in table 3 show that the intercept dummy and the three dummy interaction variables on terms containing the row index i are significantly different from zero; in contrast, the dummy interaction variables on the terms containing the column index are not significantly different from zero. The significance of these dummies indicate that the shapes of the urban and rural multiplier surfaces are significantly different.

These results also indicate that the absolute values of the coefficients on i , i^2 , and i^3 are closer to zero for the rural regions than for the urban

region. Thus, the urban surface has a steeper, more pronounced hill-shape surface, indicating that the differences in strength of linkage between middle-income households and other households are larger in the urban region than in the rural regions.

Summary and Conclusions

The results of this study demonstrate important interhousehold linkages that are ignored in standard regional input-output models. While these differences in household linkages are important when changes in exogenous household income are investigated, they will also affect the accuracy of sectoral multipliers. The results further indicate that higher-income households are not strongly linked to low-income households, at least in the four study regions. In addition, given equal changes in exogenous income, low-income households will have much higher impacts on the local economy than high-income households.

The results also indicate that more complex economies can be characterized as having much stronger linkages between households. The results indicate significant rural-urban differences in the level and shape of the multiplier surfaces but lend little support for the hypothesis that household linkages are more uniform in more developed economies than in the less developed, rural economies.

The disaggregation, by income, of the household sector in this study represents an important improvement over standard regional input-output models. While not a new technique, it has only recently been employed at the state level (Martin and Henry, Rose and Beaumont). Our study extends this work by demonstrating the feasibility of such a model for small, substate regions.

[Received January 1990; final revision received June 1990.]

Table 3. Household Multiplier Surface Equation Comparing Rural and Urban Regions

Variable	Coefficient	Estimated Parameter
Intercept	b_0	13.96**
i	b_1	-3.18**
j	b_2	-3.59**
i^2	b_3	1.01**
j^2	b_4	0.55**
ij	b_5	-0.05**
i^3	b_6	-0.06**
j^3	b_7	-0.03**
Dummy variables on:		
Constant	d_1	11.70**
i	d_2	-11.22**
i^2	d_3	2.30**
i^3	d_4	-0.12**
j	d_5	-0.59
j^2	d_6	0.09
j^3	d_7	0.05
R^2		0.57

^aAsterisk indicates significant at the 10% level; double asterisk indicates significant at the 1% level.

⁶ The same equation was also run with dummies for each of the rural regions. These region-specific variables were not significantly different from each other. The more detailed equation results are not presented.

References

- Bernat, G. Andrew, Jr. "Impacts of Transfer Payments." Paper presented at Economic Development Related Applications of Input/Output Models, Kansas City MO, 28 Feb.-2 March 1989.
- . "Income Distribution in Virginia: The Effect of Intersectoral Linkages on the Short-Run Size Distribution of Income in Small Regions." Ph.D. thesis, Virginia Polytechnic Institute and State University, 1985.
- Johnson, Thomas G., and Oral Capps, Jr. "Rural Area

- Consumer Demand and Regional Input-Output Analysis: Comment." *Amer. J. Agr. Econ.* 66(1984):173-76.
- Martin, Thomas L., and Mark S. Henry. "Rural Area Consumer Demand and Regional Input-Output Analysis." *Amer. J. Agr. Econ.* 64(1982):752-55.
- Miyazawa, K. *Input-Output Analysis and the Structure of Income Distribution*. New York: Springer-Verlag, 1976.
- Park, Se-Hark, Malek Mohtadi, and Atif Kubursi. "Errors in Regional Nonsurvey Input-Output Models: Analytical and Simulation Results." *J. Rgnl. Sci.* 21(1981): 321-39.
- Rose, Adam, and Paul Beaumont. "Interrelational Income-distribution Multipliers for the U.S. Economy." *Frontiers of Input-Output Analysis*, ed. Ronald E. Miller, Karen R. Polenske, and Adam Z. Rose, pp. 134-47. New York: Oxford University Press, 1989.
- . "Interrelational Income-distribution Multipliers for the West Virginia Economy." *J. Rgnl. Sci.* 28 (1988):461-75.
- Rose, Adam, Brandt Stevens, and Gregg Davis. *Natural Resource Policy and Income Distribution*. Baltimore MD: Johns Hopkins University Press, 1988.
- Schultz, Siegfried. "Approaches to Identifying Key Sectors Empirically by Means of Input-Output Analysis." *J. Develop. Stud.* 14(1977):75-96.
- Stevens, Benjamin H., and Glynnis A. Trainer. "Error Generation in Regional Input-Output Analysis and Its Implications for Non-Survey Models." Amherst MA: Regional Science Research Institute Disc. Pap. No. 108, 1978.
- Stevens, Benjamin H., George I. Treyz, David J. Ehrlich, and James R. Bower. "A New Technique for the Construction of Non-Survey Input-Output Models." *Int. Rgnl. Sci. Rev.* 8(1983):271-86.
- U.S. Department of Labor, Bureau of Labor Statistics. *Consumer Expenditure Survey: Diary Survey, July 1972-June 1974*. Bull. No. 1959, Washington DC, 1977a.
- . *Consumer Expenditure Survey: Interview Survey, 1972-1973*. Bull. No. 1997, Washington DC, 1977b.
- Virginia Employment Commission (VEC). *Occupational Employment Statistics: Selected Transportation, Public Utilities, and Trade Industries*. Richmond, 1981.
- Weisskoff, Richard. "Income Distribution and Export Promotion in Puerto Rico." *Advances in Input-Output Analysis*, ed. Karen R. Polenske and Jiri V. Skolka, pp. 205-28. Cambridge MA: Ballinger Publishing Co., 1976.
- Weisskoff, Richard, and Edward N. Wolff. "The Structure of Income Inequality in Puerto Rico." *J. Develop. Econ.* (1981):206-28.

Economic Impacts, Value Added, and Benefits in Regional Project Analysis

Joel R. Hamilton, Norman K. Whittlesey, M. Henry Robison, and John Ellis

This paper addresses five issues encountered when estimating secondary benefits in regional project analysis: (a) the correction for opportunity cost of factors used, (b) the treatment of mobile factors, (c) the effect of economies of size, (d) the role of forward linkages, and (e) the role of spatial structure of economic regions. The first four are reasons that only a small part, if any, of regional impacts can be treated as regional net benefits. The fifth is a reason that, when secondary benefits or damages do exist, their correct estimation can depend on the spatial structure of the affected areas.

Key words: economic impacts, net benefits, regional project analysis.

Economists are often asked to assess the benefits or damages from a past or future event on the economy of some region. Examples include assessment of regional economic consequences of alternative timber harvest levels, entry or exit of a manufacturing plant, or the construction or demise of an irrigation project. Analysts have increasingly turned to the tools of regional analysis, especially regional input-output (I-O) models, to estimate these benefits or damages. Regional I-O models are constructed to estimate linkages among sectors of the economy of a target region. In this way, an event affecting one sector can be traced through the regional economy, and the change in value added, income earned by primary factors of production, can be estimated. This approach allows estimation of both direct impacts caused by the initial change in the affected sector and secondary impacts

which result as the direct spending works its way through the economy.

Direct income impacts of a project are the factor payments: the wages, rents, and profits earned by input factors used directly by the project. Secondary impacts result when the directly affected sector buys inputs from other sectors (backward linkages) or produces outputs that become inputs for expansion of other regional industries (forward linkages). A new irrigation project will cause agriculture to buy more from backward linked fertilizer, machinery, and insurance sectors, and may allow expansion of forward-linked livestock and food-processing sectors. The directly affected sector, along with backward and forward linked sectors generate value added consisting of income earned by the input factors. Thus, the available tools, including I-O models, allow us to estimate the impacts of the project or event on regional economic activity. The impact analysis is the easy part; the step that causes problems is translating these impacts into estimates of benefits or costs.

This paper addresses several methodological issues which arise and errors that are committed when estimates of regional economic impact are used to derive estimates of regional primary and secondary project benefits. Most of the issues noted in this paper apply whether the impact estimates are based on I-O models, economic base analyses, econometric models, or the cost ac-

Joel R. Hamilton and M. Henry Robison are a professor and an assistant professor, respectively, Department of Agricultural Economics, University of Idaho. Norman K. Whittlesey and John Ellis are a professor and an assistant professor, respectively, Department of Agricultural Economics, Washington State University.

This is Idaho Agricultural Experiment Station Paper No. 9011. This research is a contribution to Western Regional Project W-178, "Water Management and Conservation in Western Irrigated Agriculture."

The authors acknowledge helpful comments on earlier versions of this paper from Walter Butcher, Rodney Jensen, Charles Howe, Robert Young, Stephen Cooke, Roger Mann, David Walker, and the anonymous *Journal* reviewers.

counting commonly used in benefit-cost analysis. Some of these issues are noted in the regional analysis literature (Stabler, Van Kooten, and Meyer; Hamilton and Gardner), but we have found no comprehensive discussion of the issues. In fact, many of the more serious misapplications of these tools are not in the mainstream professional literature but in the myriad of environmental impact statements, forest timber plans, community development analyses, and other applied impact studies. These studies often represent the attempts of regional planners, semiprofessional economists, or project promoters to apply tools and concepts learned or mislearned from professional economists.¹ This paper addresses five related methodological issues.

First, it usually is a methodological error to use project impact on value added as the measure of project regional benefits or costs. Changes in value added estimated with an I-O model consist of changes in income accruing to regional primary inputs, which are at least partly offset by the opportunity cost of diverting these resources from alternative employment elsewhere in the regional economy. What portion of regional income impact represents actual regional net benefit or net damage from the event?

Second, because all factor income is paid to people, who may or may not move in response to the event being studied, there is an ambiguity regarding what income and whose income should be counted when computing primary and secondary regional project benefits.

Third, as the level of economic activity changes in the affected industries of a region, they shift to different points on their average cost curves. How frequently do the affected industries exhibit economies or diseconomies of size sufficient to result in significant regional benefits or damages?

Fourth, because a regional I-O model traces each sector's purchases from other sectors of the regional economy, it captures only the portion of secondary impacts caused by backward linkages from the event being studied. Under what conditions are forward secondary linkages also important and how can they be estimated?

Fifth, I-O models are usually based on political boundaries such as counties or states, while the functional economic areas within which impacts occur often cross political boundaries. What are the implications of this divergence for correct estimation of regional project benefits?

Opportunity Cost of Factors Used

It takes production inputs to support the economic activity associated with a project or event, including any backward and forward linked secondary activity. These production inputs have an opportunity cost. Using the division that is both convenient and common in the literature, this section will address only the direct income impacts and those caused by backward linkages; a later section will address forward linkages.

Benefit cost analysis commonly assumes full employment of most primary factors of production, perfect resource mobility, and absence of scale economies (Howe and Easter, p. 26; Margolis). If these assumptions apply to all factors, then, even though the value added impact might be large, benefits would be zero because all factors employed with the project could have received essentially the same returns in alternative employment without the project. In other words, the income impact resulting from project-related activity is exactly offset by the opportunity cost of the resources used.²

Most applications of project analysis do make specific departures from these assumptions of full employment and complete mobility. Often one input, such as the water used by an irrigation project, is considered underutilized and immobile. Thus the opportunity cost of this resource is low. In this case, when the project's direct impacts on value added are translated into direct benefits by subtracting the opportunity cost of the inputs used, the direct net benefits consist of the residual increase in returns earned by the previously underutilized resource. Thus, the residual returns to water, after subtracting the cost of all other project inputs, is often taken as an

¹ Just and Hueth; Just, Hueth, and Schmitz; and Cooke have addressed the measurement of welfare change in a series of forward and backward linked industries impacted by some event. While their theoretical development contributes to understanding the issues presented in this paper, it is not clear that their approach is generally empirically implementable for cases such as those discussed in this paper, nor is it likely to soon replace the I-O-based approaches.

² While not the usual case, some projects do cause inputs to be used or combined in new ways. Such technological change can enhance the productivity of resources above their opportunity costs outside the project. However, productivity enhancement benefits should be attributed to a project for only the short run. In the long run, the normal course of technology adoption should establish the enhanced productivity levels as the new regional norm.

estimate of the direct benefits of a water project.³

The assumptions of full employment and complete mobility can often be applied plausibly to all inputs used in generating secondary project impacts. Thus, regardless of the size of the estimated change in value added from secondary impacts, it may be exactly offset by opportunity costs of the inputs used, leaving net secondary project benefits of zero. This posture is especially plausible from a national accounting perspective and is codified in the "Principles and Guidelines" for water resource projects (U.S. Water Resources Council), which directs that project secondary impacts must not be included in "National Economic Development" benefits.

Using the rule that net benefits equal impacts less opportunity costs, additional primary or secondary regional net project benefits may result if factor prices and factor opportunity costs diverge such that unemployed or underemployed resources persist in the region because of barriers to resource mobility. Little empirical work has addressed the level and nature of resource unemployment needed to justify including such benefits in regional project analysis. The few studies which have addressed the issue have relied more on assumption than on analysis of real data. The most prominent is the work by Haveman and Krutilla, who studied the labor and capital inputs required for water resource construction projects. They concluded that the national opportunity cost of resources used directly and indirectly by such projects ranged from 69% to 94% of market value, depending on the type and location of the project. This conclusion implies that between 6% and 31% of the income generated by these projects could be counted as net benefits.⁴

³ Large projects may result in some input and output price impacts, but Young and Haveman stress that these impacts should not be confused with benefits: "Pecuniary impacts (usually called 'secondary' or 'indirect' economic effects) are those reflected in changes in incomes or prices caused by shifts in supply or demand. Pecuniary externalities are likely to represent income distribution rather than allocative effects, and their inclusion would amount to double counting" (p. 190).

A related argument is that projects which produce consumer goods confer benefits of increased consumer surplus. Such benefits can occur in regional project analysis only if the additional production lowers local prices and a large portion of the product is consumed locally. Of course, the loss of producers surplus (value added from the depressed local prices can equal or exceed the consumer benefits which accrue to regional consumers. Such price change effects are difficult to address within the fixed price context of I-O modeling.

⁴ In the context of benefit-cost accounting, project gross benefits are unaffected by whether some of the factors used were otherwise underemployed or unemployed. However, the difference between the factor prices and their opportunity cost is a "credit" offsetting

In his *ex post* study of the Colorado-Big Thompson water project, Howe (1987) assumed that 20% of capital earnings and 23% of wages and salaries represent net benefits. Hamilton and Gardner assumed that 10% to 20% of secondary value added by new irrigated agriculture in southern Idaho could be counted as net secondary benefits. In neither case were these assumptions based on empirical analysis.

Howe and Easter argue that if a region is highly dependent on a vulnerable industry and resources are immobile, then a much larger part of the value added resulting from a "rescue" of that industry can be counted as net benefit. However, Howe's "rescue" conditions will rarely, if ever, be met in practice. Even in Howe's negative impact scenario most of the released resources, except for some undepreciated immobile fixed assets, will eventually find other employment. The lifetime of most projects will be long enough to allow resources to move to their best use and to negate cyclical unemployment as an argument for secondary benefits. Even when short-run employment benefits exist, they can rarely be justified over the entire project life.

Most cases justify conclusions such as those reached in a chapter written by Hanke and Walker in the book by Haveman and Margolis: "Since there is no evidence that the conditions put forward by Margolis exist in the prosperous Mid-State area, secondary benefits have been eliminated from the analysis because they represent pecuniary transfers and not real effects" (Hanke and Walker, p. 343). Stabler, Van Kooten, and Meyer say: "Disregarding the possibility of declining long-run average cost curves in linked activities within the region, the general case for calculating secondary benefits on the basis of employable and immobile resources in Western North America is weak" (p. 20).

The burden of proof is on those who claim that a project uses labor or capital which would be unemployed without the project, therefore justifying the existence of secondary regional project net benefits, or the inclusion of additional value-added components in primary regional net benefits. Under most circumstances, the convergence between input prices and their opportunity costs restricts project direct regional net benefits to the residual returns to some particular underutilized target resource, such as the

some of the factor payments charged in the project cost account. Thus employment generation should not be counted a project benefit, but access to otherwise underutilized resources can reduce project cost, contributing to project net benefits.

water used by an irrigation project. Similarly, only a small portion (if any) of secondary income can be counted as secondary regional net benefits.

Treatment of Mobile Resources

Changes in direct and secondary regional economic activity of the kind discussed above usually result in some interregional movement of labor and capital. All factor income, whether earned by labor, capital or land, is paid to people. However, the people who receive that factor income do not always move in the same pattern as the labor and capital resources themselves. This results in a problem in deciding whose income counts when regional project benefits or costs are computed. This problem relates to the issue of accounting stance discussed in Stabler, Van Kooten, and Meyer.

The two polar possibilities are to count all or to count none of the income earned by immigrant resources. By definition regional income is the sum of the incomes of all residents of a region. Thus, the "all" posture would compare the sum of incomes accruing to everyone who would have lived in the region without the project and sum of incomes of everyone who will be living in the region with the project—with the difference proposed as an estimate of regional project benefits. While this approach occurs frequently in the project analysis literature, we question its validity.

The "none" alternative of not counting any of the income received by immigrants was advocated by Howe:

The water manager . . . should act in the interests of the existing population of his area unless he receives policy direction from his constituency to the contrary. This point of view is the same as that taken by corporate financial managers in deciding the merits of new investments or stock issues, and is expressed in the question, 'Will the action dilute or enhance the equity of existing stockholders?' If this stockholder equity view were taken by city managements in general, continued urban growth might be seen in a different light. (page 12)

Similar positions were stated by Hamilton and Gardner:

Insofar as regional project evaluation is concerned, it is the differences in value added generated by resources that would be in the region in the absence of the project which are relevant. It is debatable whether wages and salaries going to imported labor

and returns to capital provided from outside the region (while a part of gross regional product), are valid regional benefits to those people who would have been residents of the region without the development. (page 8)

And, again, we see agreement in this statement by Keith and Glover:

There is considerable controversy concerning the payments to imported labor as a regional benefit. If workers are transient, then long-term changes in payments to labor are not likely to result in large gains to the local economy. As long as it is a permanent change, imported labor represents an increase in resident population and household income. However, these benefits do not necessarily accrue to the population which was resident in the region prior to the project development, even though a substantial part of increasing service costs (such as schools, roads, etc.) and/or other costs (housing, food, etc.) may be borne by current residents. Thus there is a choice between estimating benefits to existing or to projected populations. If regional population growth is desirable, or is a direct goal of a project, then perhaps benefit estimates should include net income to permanent immigrants in the benefit calculations. However, any increasing costs of consumer goods or public services should be netted from the calculated benefits. It is highly doubtful that payments to imported capital will be paid to residents of the local community. In order to be as conservative as possible, to avoid counting benefits which may be fugitive (i.e., may not remain in the region), and to minimize the problems of calculating increases in, and distribution of, increasing social costs, it is recommended that only benefits accruing to current residents of the region be calculated. (page 22)

Consider an example. Assume that the impact of a new project causes someone from another community to take a new job in your community. Further, assume that the migrant's previous employment income was \$30,000 per year and that person now earns \$40,000 in the new job. Note that the aggregate accounting region consisting of both donor and recipient communities together has clearly experienced a \$10,000 benefit; and \$40,000 income increase in one community offset by a \$30,000 opportunity cost in the other. However, from this aggregate accounting stance incorporating both the mover's origin and destination, no net resource migration has occurred.

The problem arises in measuring the benefit of this move to your community. Certainly the new individual's local spending for consumption or investment will stimulate local businesses, and some of this induced income may find its way to the residents of the community.

Other than this induced effect from respending, however, neither you nor other original members of your community benefit from the \$40,000 income per se. Similarly, the remaining residents of the community that lost \$30,000 employee may suffer some income loss because respending of this income has stopped, but they do not feel the loss of the \$30,000 per se. It is the migrant alone who is the direct beneficiary of the \$10,000 increase in factor payments.

Whose welfare should be included within the accounting stance used to determine the well-being of the region? To what extent does well-being of neighbors, especially new neighbors, enter into an individual's welfare function? For such questions economic theory can give little guidance. One plausible approach for a recipient community inclined to value growth is to count the \$10,000 income net of opportunity cost (but not the entire \$40,000 income of the migrant) as a net benefit. This is the benefit to the community residents from a happy new resident who has just experienced a \$10,000 income increase. For the donor community, the emigrant may be out of sight and out of mind. Realistically, migrants are neither completely included in nor completely excluded from the community accounting perspective. The accounting stance advocated by Young and Howe, Hamilton and Gardner, and Keith and Glover would eliminate all direct benefits consisting of immigrant wages and count only benefits accruing to previous residents.

It is important to examine the assumptions implicit in this example. First, a \$10,000 increment between donor and recipient regions could not coexist with full employment and perfect factor mobility. Second, if such a differential between the two regions did exist, it could only persist in the short run, supported by a labor immobility that the migrant has managed to overcome. In the longer run, either wages in the donor region will be bid up by continued out-migration, or wages in the recipient region will be depressed by continued immigration. Thus, the benefits to the recipient region attributable to migrant wage changes should be either non-existent or short run.

What about the induced effects from the respending of the migrant's \$40,000 income in the destination region? Clearly, as noted in the previous section, the multiplier effects will generate further secondary regional income. However, this increase in economic activity will draw resources away from alternative employment, so the opportunity costs of these resources must

be deducted. Moreover, the induced effect will very likely be supported by the immigration of additional labor and capital, aggravating the problem of deciding whose income to count.

The situation is clearer for transient migrant labor. Income earned by labor that resides only temporarily in a region, while by definition a part of regional income, is certainly not valued very highly by other regional residents except as its respending induces activity elsewhere in the regional economy.

Similar problems are encountered with mobile capital. A new project will have both primary and secondary impacts on regional capital use and income. If internally owned capital is used, it has an opportunity cost equal to its value in its best alternative use. Inflows of monetary capital are not per se a regional benefit. It is the investment of this money in physical assets and the subsequent productive use of these assets that will generate regional incomes. However, capital attracted from outside the region, does not have zero cost, as implied in Howe (p. 79): "Insofar as these productive factors, including capital, are attracted from outside the project region, the loss of their outputs elsewhere will not be counted as costs in the project region's assessment of the project." Howe fails to note that importing this capital probably does not change its ownership. The price of importing capital to a region is the outflows of interest payments and profits sufficient to attract the capital, which is its exact monetary cost in the importing region and should approach its opportunity cost in the exporting region. Even for projects funded by government grant, this capital may have an opportunity cost in terms of other regional projects not funded (congressional logrolling also has its limits).

If a project attracts both capital and its owners as immigrants to the region, this clearly has a positive impact on regional income. The situation is similar to that for immigrant labor—one can question whether increases in profit and interest payment accruing to immigrant capitalists are per se a regional benefit. The existence of regional benefits depends on whether this new profit and interest income satisfy growth goals of the original community residents, and whether they can capture benefits from the economic activity induced by the local respending by the immigrant capitalists.

Another aspect of migrant resources is the impact that immigrant resources can have on local production processes. Immigration of labor or capital might enhance the productivity and re-

turns earned by local labor and capital. These are potentially important project benefits, which unfortunately are poorly modeled by fixed coefficient approaches such as I-O models. More research could productively be directed toward developing proper methods of incorporating these resource productivity benefits into project analysis.

Because much of the impact of projects is on income earned by immigrant labor and imported capital, this is one further reason the portion of estimated income impact, both primary and secondary, that can be considered a regional net benefit ranges from small to none.

Economies of Size

Reference was made above to changes in industry size as a justification for additional project benefits. In one sense, this is another way of viewing the underemployment/unemployment issues discussed above. Established firms in the region (the grocery store, the motel, the fertilizer dealer, the public school, and even the municipal government) may have adequate physical plant and labor to handle some increase in business. Similarly, on the downside, some decline in business may be accommodated without exit of labor or capital. Firms with considerable fixed labor or capital almost surely exhibit short-run size economies. On the other hand, if a firm's labor and capital can accommodate an expansion, then they must previously have been underemployed, or if the firm can tolerate a downturn without releasing labor or capital, that labor and capital must then become underemployed. In the long run, fixed assets must be replaced, and labor will move, ending this kind of short-run size economy benefits.

The issue goes beyond the fixed input/underemployment relationship. Long-run size economies may mean that a larger firm is better able to utilize labor or capital or has better market access than a small firm. Expansion of an industry may also result in agglomeration economies. Thus, an irrigation project might make possible the expansion of a livestock industry sufficient to give the meat-processing industry a cost advantage over competing regions. Regional growth may allow establishment of larger size, lower cost grocery stores, giving consumers the benefit of lower cost food.

On the other hand, any benefits flowing from industry size economies may be offset by other affected industries with either long- or short-run

size diseconomies. Power generation in the Pacific Northwest experiences an increase in average cost as demand grows; present generation is based on cheap hydropower methods, while increases will have to come from expensive thermal sources. Expanding firms that need to add or replace fixed assets may operate only at costs above those of an established firm with depreciated fixed assets.

Because the I-O model is based on fixed coefficients, it is of little use for estimating size economy impacts. Attempts to estimate these impacts must be based on calculations external to the I-O model. Little empirical research has documented the role of economies of size in the affected industries as a determinant of the regional benefits or costs of resource development projects. Among the exceptions are work by Whittlesey et al., and Findeis and Whittlesey. They considered the incremental costs of providing energy supplies, incremental farm capital needs, and incremental social overhead costs associated with irrigation expansion.

Economies of size is a possible source for both primary and secondary project net benefits. However, because the fixed-coefficient I-O model is unable to incorporate size economies, it cannot be used to estimate these benefits. Further, the possible existence of size economies is not a valid justification for including I-O based estimates of secondary impacts as benefits. This is another area where research could be fruitful.

Forward Linkages

Forward linkage impacts occur if output from the directly affected sector serves as input to some other sector, thus changing the output level of that sector. For example, a new irrigation project may increase feed grain production, promoting an expansion in the regional livestock industry, which in turn could result in more meat processing in the region. As with backward linkages, some of the regional secondary income impacts produced by forward linkages may be valid components of regional secondary net benefits.

The I-O model, because it traces purchases of each sector from all others, is well suited for translating a primary impact on one sector into estimates of the secondary impacts on production and value added by the backward linked sectors of the regional economy. Backward linkages are sure to exist—production requires inputs, so changes in production will be felt by

sectors that supply those inputs. Forward linkage impacts are far less certain. If a region produces more feed grain, there is no assurance that regional livestock production will expand. If a region produces more logs, there is no assurance that they will be processed locally into dimension lumber, plywood, or houses.

Regional forward linkages exist only if the forward industries are otherwise constrained by input shortage and their expansion is a viable, profitable proposition. Often, the expansion potential of forward linked industries depends more on the regional and national market conditions for their outputs (meat, houses) than on the regional supply of raw inputs (feed grains, logs). If circumstances are unfavorable for the expansion of forward linked sectors, several possible scenarios exist. First, the raw product may simply be exported from the region. Second, especially for products that are hard to transport long distances, the increase in raw material production in the project area may cause regional declines in product prices, leading to offsetting production declines inside or outside the accounting region. Third, forward linked activity using inputs produced in the project area might increase, but an offsetting demise of such processing activities could occur elsewhere in or out of the region. For the first two scenarios, no regional forward linkage impact would occur. For the third, any regional impact occurs through transfer from another region.

The weakness of forward linkages is recognized by Hamilton and Gardner:

Certainly some of the output of newly developed land will be fed to livestock, other crops from this land will be used as input to food processing plants, and these secondary activities will generate secondary value added. However, the demand for regional exports of processed food and livestock products is quite inelastic and driven largely by factors outside the region. We question the extent to which new irrigation actually increases the total amount of secondary food processing or livestock production in a region. Development may only mean that new farms rather than the farms present without development will supply a share of the inputs to these secondary activities.

(page 8)

Due to inelasticities of crop demand, the net effect of new irrigation on state crop acreage has historically been an increase in small grain and forage crop acreage with little increase in the acreage of higher valued crops. Essentially the acreage of high valued crops displaces existing acreage of these crops. (This treatment conforms to the federal Principles and Guidelines which mandate that returns to high

valued specialty crops not be counted as part for the direct benefits of a water resource project.)
(page 5)

This same weakness is recognized by Howe:

There is no question that the availability of project agricultural outputs and the related demands for inputs led to the expansion of both forward (output processing) and backward (input supplying) linked activities in the project region. The sectors that are strongly forward linked to agriculture in the region are livestock and food processing. The markets for these final products are found almost completely outside the C-BT project region. It seems reasonable to argue that, had C-BT not existed, the additional supplies of these final products would have been produced elsewhere in the western United States. For this reason forward linkages can be ignored when computing net income changes from the national accounting stance, although not from the project region stance.

(page 88)

If a project does cause secondary forward linkage impacts these will be in the form of output changes in other sectors and income earned by the production inputs used. The amount of this income impact that counts as regional benefit depends on the same opportunity cost considerations outlined for backward linkages. Labor and capital income impacts generated by forward linkages count as benefits only to the extent that the return in the new use exceeds its opportunity cost in its alternative use, and if some of these returns are captured by the permanent residents of the region. This will only be true if unemployment and immobility of labor and capital are substantial.

Young and Howe outline steps necessary to properly incorporate forward linkages into analysis of regional project benefits in instances when this is justified by market conditions and profitability of prospective industries. To estimate regional forward linkages if sector A increases output of raw material by Q_A , they note that one should:

- a. check the structure of local industries to see whether or not forward linkages appear likely;
- b. calculate the added processing output, Q_P , that is likely to follow from the availability of Q_A ;
- c. insert Q_P in the input-output (inverse) model to determine what output changes will be required from all sectors other than that containing A;
- d. calculate the increased factor payments (i.e., to labor (households), rent, interest, profits) that the model predicts to follow from the increased output in all sectors other than A;
- e. estimate the opportunity costs of these newly employed factors to the state and subtract them from

the increased payments calculated in (d) to get secondary net benefits. (page 69)

The issue is whether forward linkages exist, whether they fall within the accounting stance, and how they should be treated in empirical project analyses. We have outlined conceptual reasons that forward linkage benefits may be nonexistent or small. We conclude that the existence of forward linkage benefits must be justified empirically, case by case, using procedures like those outlined by Young and Howe.

Spatial Diffusion of Secondary Impacts

The above discussion has taken the project region as an internally undifferentiated partition of a uniform landscape. However, in the well-known sense of central place theory (Christaller, Losch), economic regions exist quite apart from the political regions superimposed on them, and these economic regions have a hierarchical internal and external trade structure. Regions used for input-output modeling are typically based on political units (states, counties, multistate or multicounty aggregates), easing data and model-building problems, and because the questions asked are usually framed in terms of political units. Both the trade patterns and the divergence between the boundaries of political and economic regions have consequences for regional project analysis.

Seninger noted that there may be clear policy and planning implications from an analysis of spatially diffused secondary impacts among a system of communities. Disaggregating impacts of a project to the community level may be helpful for understanding the nature of the impacts, planning how to deal with them, and deciding who should pay for them. By integrating central place and I-O theory, through an intercommunity I-O model (as in Robison and Miller), it is possible to trace the diffusion of secondary impacts through the central place hierarchy within a study region.

Ideally, I-O analysis should be used to model the economy of one or a hierarchical aggregation of functional economic areas (Richardson).⁵ When an I-O model is shoehorned to fit

the political units of data availability or to match the political boundaries of policy issues, the common nonsurvey techniques used to build these models can result in biased estimates of secondary project impacts (Robison and Miller). Economic impacts do not stop at political boundaries but usually spill across these boundaries into contiguous parts of a transcending economic region. The concept of interregional spillover is another way of looking at economic leakages from a region. Generally, small and less self-sufficient regions will have lower output and income multipliers because more spending leaks or spills into adjacent regions.

One can get a feeling for the spillover relationship from the output multipliers presented in table 1, which are based on the nested set of consistent I-O models from Hamilton and Jensen. As expected, Australia had the largest multiplier. The state of Queensland had a smaller multiplier because some secondary activity spills across its borders into other states. We can assume that each dollar of primary impact from some event in Queensland is associated with \$0.71 of secondary impacts, the same ratio as for events elsewhere in Australia. However, Queensland captures only \$0.55 of these secondary impacts while \$0.16 of secondary impacts spill over to other states. The ratio of spillover secondary effects to those captured by the region is the spillover coefficient, $0.16/0.55 = 0.291$ for Queensland. Moreton Region, containing Brisbane, the state capital, strongly dominates the economy of Queensland, allowing little of the secondary impacts from an event in that region to spill into other regions of the state. In contrast, an event in Toowoomba, a smaller town in Moreton Region, generates much more spillover secondary impacts than local secondary effects because it is so small and so strongly dominated by metropolitan Brisbane.

Interregional or multiregional I-O models (Miller and Blair) are appropriate analytic tools in cases where spillover effects likely are im-

Table 1. Australia Interregional Spillover Coefficients

Region	Average Output Multiplier	Spillover Coefficient to Next Higher-Order Region
Australia	1.71	
State of Queensland	1.55	0.291
Moreton Region	1.54	0.019
Toowoomba City	1.10	4.400

⁵ In the 1970s, the U.S. Department of Commerce, Bureau of Economic Analysis (1975, 1982) relied on hierarchical trade and central place principles to map the economic boundaries of the United States. The BEA aimed at delineating a complete set of functional economic areas for the United States (see also Fox and Kumar). BEA Economic Areas generally have a "Standard Metropolitan Statistical Area" at the center, surrounded by a number of hinterland counties. While the BEA areas follow county boundaries, they often cross state boundaries to define a region comprising parts of two or more states.

portant. This framework explicitly models both the structure of each regional economy and the linkages between regions. In this way, the impact of an event in one region can be traced to all other regions. Such a multiregional model was constructed by Carter and Ireri as part of a study of the impact of California-Arizona water reallocation alternatives. Their results, shown in table 2, can be interpreted in terms of spillover coefficients. These results indicate that for each dollar of secondary impact generated in Arizona, an additional spillover of secondary impacts ranging from \$0.058 to \$0.822 will also occur in California.

The consequences of interregional spillovers can be important in regional project analysis. Miller examined the possibility that spillovers could be the basis of interregional feedback, where development in one region can spill over, stimulating development outside the accounting region, which in turn spills back to further stimulate economic activity in the first region. If such feedbacks are significant, an I-O model would underestimate the extent of project impacts.

The recent Pecos River case in the U.S. Supreme Court (482 U.S. 124, 1987) illustrates another kind of spillover relationship based on a resource allocation conflict between two ad-

jacent political regions with an interlinked (nested) economic structure. In this case, if the accounting stance were based solely on the political boundaries, ignoring the hierarchical economic connections between the two areas, incorrect policy conclusions could easily result. The damages suffered by Texas for past underdeliveries of water by New Mexico under the Pecos River Interstate Compact include direct and secondary income losses net of the opportunity cost of the resources used. However, New Mexico would have had to reduce its irrigated acreage by at least 20 acres for each additional acre that could be irrigated in Texas. The 20 to 1 ratio is a consequence of the lags in the groundwater hydrology of the region, the evaporation losses, and the salinity buildup of Pecos River water delivered to Texas. The economies of the affected part of New Mexico are closely tied through strong forward and backward linkages to the adjacent parts of Texas. The question is: What would have been the net benefits to Texas if it had received the water and the direct and secondary impacts which that implies, but at the cost of losing spillover benefits because of the associated reduction in New Mexico irrigated acreage?

It is interesting to note that the two New Mexico counties which would have lost irrigated acreage under compact compliance are designated by BEA as part of the El Paso, Texas, functional economic area. The Texas-New Mexico border runs neatly between these New Mexico counties and the El Paso, Texas, SMSA that is the principal central place in this functional economic area. Because of the obviously close economic ties between these portions of New Mexico and Texas, actions taken in New Mexico should have large spillover effects in Texas. In this case, there is a real possibility that the spillover benefits accruing to Texas from using the water in New Mexico would exceed the direct and secondary damages of not having the water to use in Texas, leaving Texas better off because it did not get the water. However, if analysis was based on separate I-O models of the affected New Mexico and Texas areas, ignoring the economic linkages between the two, these potentially important spillover effects would remain hidden.

Note the internal consistency; if factor opportunity cost were 100% of factor market cost, then there were no secondary benefits to New Mexico from using the water and no spillover benefits to Texas; however, if factor unemployment or immobility is invoked to justify sec-

Table 2. Secondary Impact Spillover Coefficients from Arizona into California

Sector	Coefficient
Meat animals & products	0.058
Poultry & eggs	0.822
Farm dairy products	0.122
Food & feed grains	0.247
Cotton	0.102
Vegetables	0.197
Fruit & nuts not citrus	0.197
Citrus	0.145
Forage	0.199
Miscellaneous agriculture	0.102
Grain mill products	0.201
Meat, poultry process	0.124
Dairy products	0.064
Can, preserve, freeze	0.813
Misc. agr. processing	0.301
Chemicals & fertilizer	0.309
Petroleum	0.152
Fab metal & machinery	0.335
Aircraft	0.181
Primary metals	0.134
Other manufacturing	0.144
Mining	0.079
Utilities	0.062
Selected services	0.115
Trade & transport	0.058
Other services	0.115

ondary benefits, then spillovers are also a possibility.

The divergence between the political regions often used for regional modeling and the functional economic areas in which economic activity actually takes place sometimes causes analysts to incorrectly specify their accounting stance, resulting in biased estimates of project impacts and project benefits and missing important benefit spillovers between accounting regions.

Conclusions

This paper has addressed five issues that are encountered when estimating secondary benefits in regional project analysis: (a) the proper correction for the opportunity cost of factors used, (b) how to treat mobile factors, (c) the effect of economies of size, (d) the role of forward linkages, and (e) the role of the spatial structure of economic regions. The first four issues serve as reasons only a small part, if any, of regional impacts can be treated as regional net benefits or costs. The fifth is a reason that, when secondary benefits or damages do exist, their correct estimation can depend on recognition of the spatial structure of the affected areas. All five serve as cautions to the project analysis practitioner and as challenges for future conceptual and empirical research.

[Received January 1990; final revision received June 1990.]

References

- Carter, H. O., and D. Ireri. "Linkage of California-Arizona Input-Output Models to Analyze Water Transfer Patterns." Paper presented at the Fourth International Input-Output Conference, Geneva, Switzerland, Jan. 1968.
- Christaller, W. *Central Places in Southern Germany* (trans. C. W. Baskins). Englewood Cliffs NJ: Prentice-Hall, 1966.
- Cooke, S. C. *A Theory of Consumer Surplus and Economic Rent and an Application for Measuring the Benefits from the Mechanical Cucumber Harvester*. Ph.D. thesis, Michigan State University, 1985.
- Findeis, J. L., and N. K. Whittlesey. "The Secondary Economic Impacts of Irrigation Development in Washington." *West. J. Agr. Econ.* 9(1984):233-43.
- Fox, K. A., and T. K. Kumar. "The Functional Economic Area: Delineation and Implications for Economic Analysis and Policy." *Papers of the Regional Science Association*, 1965.
- Hamilton, J. R., and R. C. Jenson. "Summary Measures of Interconnectedness for Input-Output Models." *Environ. and Planning A* 15(1983):55-65.
- Hamilton, J. R., and R. L. Gardner. "Value Added and Secondary Benefits in Regional Project Evaluation: Irrigation Development in the Snake River Basin." *Ann. Rgnl. Sci.* 15(1986):1-11.
- Haveman, R. H., and J. V. Krutilla. *Unemployment, Idle Capacity, and the Evaluation of Public Expenditures: National and Regional Analysis* Baltimore: Johns Hopkins University Press, 1968.
- Howe, C. W. *Benefit-Cost Analysis for Water System Planning*. Washington, DC: American Geophysical Union, 1971.
- . "Project Benefits and Costs from National and Regional Viewpoints: Methodological Issues and Case Study of the Colorado-Big Thompson Project." *Nat. Resour. J.* 27(1987):5-20.
- Howe, C. W., and K. W. Easter. *Interbasin Transfer of Water: Economic Issues and Impacts*. Baltimore MD: Johns Hopkins Press for Resources for the Future, 1971.
- Just, R. E., and D. L. Hueth. "Welfare Measures in a Multimarket Framework." *Amer. Econ. Rev.* 69(1979):947-54.
- Just, R. E., D. L. Hueth, and A. Schmitz. *Applied Welfare Economics and Public Policy*. Englewood Cliffs NJ: Prentice-Hall, 1982.
- Keith, J. E., and T. F. Glover. "Secondary Impacts and Benefits of Water Reallocation in the Snake River Basin of Idaho." Report to the Snake River Studies Advisory Committee, Idaho Water Resources Research Institute, Moscow, Oct. 1988.
- Losch, August. *The Economics of Location*, trans. W. H. Woglom and W. F. Stolper. New Haven CT: Yale University Press, 1954.
- Margolis, J. "Secondary Benefits, External Economies, and the Justification of Public Investment." *Rev. Econ. and Statist.* 39(1957):284-91.
- Miller, R. E. "Interregional Feedback Effects in Input-Output Models: Some Preliminary Results." *Pap. Rgnl. Sci. Assoc.* 17(1966):105-125.
- Miller, R. E. and P. D. Blair. *Input-Output Analysis: Foundations and Extensions*. Englewood Cliffs NJ: Prentice-Hall, 1985.
- Richardson, H. W. *Input-Output and Regional Economics*, New York: Halsted Press, 1972.
- Robison, M. H., and J. R. Miller. "Cross-hauling and Nonsurvey Input-Output Models: Some Lessons from Small-Area Timber Economies." *Environ. and Planning A* 20(1988):1523-30.
- Seninger, S. F. "Expenditure Diffusion in Central Place Hierarchies: Regional Policy and Planning Aspects." *J. Rgnl. Sci.* 18(1978):243-61.
- Stabler, J. C., G. C. Van Kooten, and N. Meyer. "Methodological Issues in the Evaluation of Regional Resource Development Projects." *Ann. Rgnl. Sci.* 22(1988):13-25.
- U.S. Department of Commerce, Bureau of Economic Analysis, Regional Economic Analysis Division. *1980 OBERS BEA Regional Projections: Economic Activity*

- in the United States by State, Economic Area, SMSA and State Portions of the Areas, Historical and Projected—1969–2030. Washington DC, July 1982.
- . "The BEA Economic Areas: Structural Changes and Growth, 1950–73." *Survey Current Bus.*, Nov. 1975.
- U.S. Water Resources Council. "Economic and Environmental Principles and Guidelines for Water and Related Land Resources Implementation Studies." *Federal Register*, 10 March 1983.
- Whittlesey, N. K., Gobind S. Bhagia, Leroy Blakeslee, Walter R. Butcher, Kenneth C. Gibbs, William Gray, W. Smith Greig, and Donal A. West. *Benefits and Costs of Irrigation Development in Washington, Vol. II: Final Report*. Dep. Agr. Econ. Washington State University, Oct. 1976.
- Young, R. A. and R. H. Haveman. "Economics of Water Resources: A Survey." *Handbook of Natural Resource and Energy Economics (Vol. II)*, ed. A. V. Kneese and J. L. Sweeney, chap. 11. Amsterdam: North-Holland Publishing Co., 1985.
- Young, R. A., and C. W. Howe. "Handbook for the Economic Evaluation of Application for Appropriation of Surface and Groundwater in the State of Idaho." Report to the Snake River Studies Advisory Committee, Water Resources Research Institute, Moscow, Idaho, Feb. 1988.

Technical Change, Land Quality, and Income Distribution: A General Equilibrium Analysis

Ian A. Coxhead and Peter G. Warr

This paper investigates the distributional effects of technical progress in agriculture. A small computable general equilibrium model is developed in which agricultural production occurs in two environments with fixed land resources of different quality. The analytical structure permits not only the rate but also the factor biases of technical change to differ between environments. Distributional results are generated for several classes of households distinguished both by their ownership of factors and by their consumption patterns. The results demonstrate that technical change produces distributional outcomes which are highly sensitive to its factor biases and to prior differences in land quality.

Key words: general equilibrium, income distribution, Philippines, technical change.

The quality of agricultural land, particularly access to irrigation, is a major factor influencing the rate of return to adoption of new technologies by farmers in developing countries. The development of new biological technologies, such as high-yielding varieties of rice, has focused on irrigated environments, where the likelihood of substantial yield increases has been highest. Accordingly, in most Asian countries, the yield gap of staple crops like rice between irrigated and less favorable environments has widened over the last two decades (Lipton and Longhurst).

Virtually since the beginning of the green revolution, concern has been expressed about the income distributional implications of the observed unevenness in the rate of growth of agricultural land productivity. The possibility that new technologies might adversely affect income distribution arises because irrigation systems are typically large and costly engineering projects. In any length of run shorter than that required for their construction, agricultural land is specific to a particular environment, irrigated or

unirrigated. The microeconomic theory of general equilibrium changes in the relative returns to sector-specific factors under a variety of exogenous shocks is well developed (Jones, Jones and Easton). The specificity of land has been exploited in a number of partial equilibrium analyses of technical change and income distribution (Evenson, Quizon and Binswanger 1983). General equilibrium models of agricultural development continue, however, to treat agricultural land as homogenous and farmers as equally affected by technical progress (Bautista, Quizon and Binswanger 1986).

This paper investigates the income distributional effects of technical progress in the agricultural sector of a small open economy. A small general equilibrium model is developed in which agricultural production takes place in two environments characterized by fixed land resources, and which accordingly experience differential rates of technical progress even while facing the same prices for output and variable inputs. The analytical structure permits not only the rate but also the factor biases of technical change to differ between environments. Distributional results are generated for several classes of households distinguished by their ownership of factors and by their patterns of consumption.

The analysis distinguishes between internationally traded goods (tradables), whose prices are determined in world markets, and nontraded goods (nontradables), whose prices are estab-

The authors are, respectively, an assistant professor, Department of Agricultural Economics, University of Wisconsin, and John Crawford Professor of Agricultural Economics, Research School of Pacific Studies, Australian National University.

Preparation of this paper was supported by funding from the Australian Centre for International Agricultural Research, Project No. 8838.

The paper has benefitted from the helpful comments of two anonymous *Journal* referees.

lished entirely in domestic markets. Neutral technical change in a sector producing tradables has effects equivalent to those of an exogenous rise in its output price. Subsequent changes in endogenous factor and commodity prices can be divided into those resulting from the movements of domestic resources required to restore equilibrium in factor markets at a given level of consumer spending, called the resource movement effect, and those resulting from the increase in consumer spending made possible by technical progress, the spending effect (Corden and Neary).

The extent of relative factor price changes, and therefore of changes in the functional income distribution, is determined by the parameters governing factor demand and supply in each sector. Changes in the relative incomes of households (the owners of factors) are determined by the distribution of factor ownership and by the composition of household consumption bundles.

Two important results emerge from the analysis. The first is that following a technical change "shock," general equilibrium feedback from nonagricultural sectors of the economy will produce changes in income distribution. Such changes are not captured in partial equilibrium models of the agricultural sector on its own. The second is that the assumption, common to many general equilibrium models, that technical change is neutral with respect to factors potentially conceals its redistributive effects among groups of households with unequal endowments of land, labor, and other productive assets.

An illustration with Philippine data provides a convenient vehicle for discussion of these results. The parameters of this numerical form of the model are derived from the available empirical research on the Philippine economy. Nevertheless, the model is deliberately stylized and simplified so as to draw out the central analytical issues. Thus, the numerical simulations are intended primarily to serve this analytical purpose.

The Model and Solution Procedure

The model belongs to the Johansen class of general equilibrium models. All models of this type are linear in percentage changes of variables. This imbues them with great flexibility in the evaluation of comparative statics results from exogenous shocks. One advantage of particular relevance is that changes in variables deriving from two or more separate shocks are simply the

sums of changes due to individual shocks. In the simulations, this linearity property of the model is exploited to decompose changes in endogenous variables. The components of these changes comprise adjustments in individual factor and commodity markets in response to factor-neutral and factor-biased technical change shocks.

The model describes a small open economy in which three commodities are produced in four sectors. Two sectors (labeled 1 and 2) produce a single agricultural good, while sector 3 produces services, and sector 4 manufactures. Sector 3 output is the only product which is not traded internationally at given world prices. It is thus the only output price to be determined within the model. Each sector employs the mobile factors, labor and capital, and a specific factor. The specific factors in the agricultural sectors are endowments of land: irrigated land in sector 1 and unirrigated land in sector 2. The specific factors in services and manufacturing represent plant, buildings, and other "bolted down" capital.

The analytical structure is a generalized form of the model analyzed in Cassing and Warr. To capture the essence of the problem, the number of sectors has been increased from the three described by Cassing and Warr to the four defined above, and each sector is endowed with a specific factor. The higher dimensionality makes theoretical analysis of the model cumbersome and the focus, therefore, is on numerical simulations. Readers wishing a more detailed discussion of the analytical properties of this class of models may refer to Cassing and Warr.

The equations, variables, and parameters of the model are set out in tables 1 and 2. The equations fall into four groups. The first two describe, in turn, factor demand and product supply, and the supply of mobile factors. Firms employ factors from the aggregate pools of mobile labor and capital as well as using sector-specific factors, the supplies of which are exogenously determined. The third set of equations describes the activities of households as earners and as consumers. Each household contributes to the aggregate supply of factors and receives income commensurate with the returns on those assets. Household incomes are exhausted in the purchase of goods. The equations in this group describe, in turn, changes in household income, household-specific price indices, real household incomes (deflated by their price indices), household consumer demands, and aggregate consumer demand.

The fourth group of equations determines the prices of factors and of the nontradable com-

Table 1. Equations in the Model

Expression ^a	Number of Equations
1. Factor demand and product supply	
$L'_s = \beta_{il}w' + \beta_{lk}r' + \beta_{ly}P'_s + \beta_{iz}Z'_s + E'_{ls}$	(s)
$K'_s = \beta_{kl}w' + \beta_{kk}r' + \beta_{ky}P'_s + \beta_{kz}Z'_s + E'_{ks}$	(s)
$Y'_s = \beta_{yl}w' + \beta_{yk}r' + \beta_{yy}P'_s + \beta_{yz}Z'_s + E'_{ys}$	(s)
2. Factor supply	
$L' = \varepsilon_l w' + \bar{L}'$	(1)
$K' = \varepsilon_k r' + \bar{K}'$	(1)
3. Household income and expenditure	
$M'_h = \delta_{hl}(1 + \varepsilon_{hl})w' + \delta_{hk}(1 + \varepsilon_{hk})r' + \sum_{s=1}^4 \gamma_{hs}(z'_s + Z'_s) + \delta_{hl}\bar{L}' + \delta_{hk}\bar{K}'$	(h)
$P'_h = \mu_{h3}P'_3$	(h)
$R'_h = M'_h - P'_h$	(h)
$C'_{h3} = \xi_{h3}P'_3 + \eta_{h3}M'_h$	(h)
$C'_3 = \sum_h \psi_h C'_{h3}$	(1)
4. Price setting and market clearing	
$0 = \alpha_{ls}w' + \alpha_{ks}r' + \alpha_{zs}z'_s - T'_s \quad (s = 1, 2)$	{ (s)
$P'_3 = \alpha_{l3}w' + \alpha_{k3}r' + \alpha_{z3}z'_3 \quad (s = 3)$	
$0 = \alpha_{l4}w' + \alpha_{k4}r' + \alpha_{z4}z'_4 \quad (s = 4)$	
$Y'_3 - C'_3 = 0$	
$L' - \sum_s \lambda_{ls}L'_s = 0$	(1)
$K' - \sum_s \lambda_{ks}K'_s = 0$	(1)
Total number of equations	4s + 4h + 6

^a Prime superscripts indicate proportional changes. Thus $X' = dX/X$. Variable and parameter definitions are listed in table 2. Sectors are indexed by s , households by h . The subscripts l , k , and y denote labor, capital and output, respectively. Numerical subscripts refer to individual sectors.

modity. It comprises four sectoral zero pure profit conditions based on the assumption of constant returns to scale and three market-clearing conditions: one for the endogenously priced nontradable (services), and one for each of the domestically mobile, endogenously priced factors, labor and capital. These seven equations simultaneously determine changes in the model's seven endogenous prices: the price of services, the prices of labor and capital, and returns to the four specific factors.

The market-clearing conditions for final commodities require explanation. The model implies two market-clearing conditions for final commodities: one for the nontradable stating that changes in the demand for and supply of that commodity must be equal and one for tradables stating that changes in the value at international prices of exports must be matched by changes in the value of imports, implying that the level

of the trade deficit or surplus is held constant. Of course, there are two traded commodities, but only one market-clearing condition, the trade balance constraint, linking changes in the excess supply of one (exports) to changes in the excess demand for the other (imports). Walras' law now implies that provided the model's consumers operate on their budget constraints, one of the two market-clearing conditions is redundant. Either one implies the other. Following Cassing and Warr, it is convenient to retain the market-clearing condition for nontradables and to suppress the trade balance constraint.¹ This choice is purely a matter of convenience; retaining the trade balance constraint and suppressing the market-clearing condition for nontradables would have given exactly equivalent results.²

Because variables enter the model only in percentage change form, it is not necessary to specify their initial absolute magnitudes or to calibrate solution values by assigning a real-valued intercept to each equation. Only the relative magnitudes of variables are used, expressed as sectoral shares in GDP and factor demand, sectoral distributive shares of factors, and household shares in consumption and factor supply. Parameters of the equations describing factor demand and product supply are elasticities. For the two agricultural sectors, these are taken from the econometric estimates reported in Coxhead. All data are contained in a working paper available upon request from the authors.

Technical Change

Three hypothetical forms of technical change are simulated in the analysis: (a) factor neutral; (b) labor saving, capital using; and (c) labor using, land saving. The nature and magnitudes of these

¹ This specification of the general equilibrium system does not explicitly include the demand functions for the two tradable commodities. The reason is that the relevant parameters are implied in full by the demand function for nontradables and the assumptions of the model. This arises because: (a) the fixed relative prices for the two tradable commodities make them a single Hicksian composite good from the consumers' standpoint, and so, effectively, only two goods are consumed—the composite tradable and the nontradable; and (b) the usual assumptions that demand functions are homogenous of degree zero and that consumers satisfy their budget constraints exactly now enable all the usual elasticity properties of the demand function for tradables to be inferred directly from the demand function for nontradables, using the familiar aggregation conditions.

² Generalizing this discussion, if there are n tradables and m nontradables, there are $m + 1$ market-clearing conditions: one for each nontradable and a trade balance constraint. One is redundant. Any m of these equations determine the m -endogenous relative prices in the model.

Table 2. Variables and Parameters in the Model

Symbol	Definition	Number of Equations
Endogenous Variables		
L_s	Labor demand in sector s	(s)
K_s	Capital demand in sector s	(s)
Y_s	Product supply in sector s	(s)
L	Aggregate labor supply	(1)
K	Aggregate capital supply	(1)
M_h	Income of household group h	(h)
P_h	Expenditure share-weighted price index of household group h	(h)
R_h	Real income of household group h (M_h/P_h)	(h)
C_{h3}	Demand for good 3 by household group h	(h)
C_3	Aggregate demand for good 3	(1)
w	Price of labor	(1)
r	Price of capital	(1)
z_s	Return to specific factor in sector s	(s)
P_3	Price of nontraded good (good 3)	(1)
	Total endogenous variables	$4s + 4h + 6$
Exogenous variables		
Z_s	Endowment of fixed factor specific to sector s	(s)
\bar{L}	Aggregate labor endowment	(1)
\bar{K}	Aggregate capital endowment	(1)
E_{is}	Technical change shifter ^a for factor i in agricultural sector s	(4)
E_{ys}	Technical change shifter ^a for output in agricultural sector s	(2)
T_s	Overall rate of technical change ^a in agricultural sector s	(2)
	Total exogenous variables	$s + 11$
	Total number of variables	$5s + 4h + 17$
Numéraire price		
P_1	Price of agricultural good	
Parameters		
α_{is}	Share of factor i in total costs of production in sector s	
β_{ij}^s	Elasticity of demand for factor i with respect to factor price j in sector s	
β_{iy}^s	Elasticity of demand for factor i with respect to output price y in sector s	
β_{iz}^s	Elasticity of demand for factor i with respect to fixed factor z in sector s	
β_{yi}^s	Elasticity of supply of good y with respect to factor price i in sector s	
β_{yy}^s	Elasticity of supply of good y with respect to own price in sector s	
β_{yz}^s	Elasticity of supply of good y with respect to specific factor z in sector s	
λ_{is}	Employment share of factor i in sector s	
ε_{hi}	Own-price elasticity of supply of factor i from household group h	
ε_i	Aggregate own-price supply elasticity of factor i	
δ_{hi}	Share of income of household group h derived from earnings of mobile factor i	
γ_{hs}	Share of income of household group h derived from earnings of specific factor Z_s	
μ_{h3}	Expenditure share of household group h on good 3	
η_{h3}	Expenditure elasticity of demand for good 3 by household group h	
ϕ_{hi}	Share of household group h in ownership of factor i	
ψ_{h3}	Share of household group h in consumer demand for good 3	
ξ_{h3}	Price elasticity of demand for good 3 by household group h	

^a For definitions, see text.

shocks are illustrative of individual technical innovations in developing country agriculture. Each type of shock, or any combination, may be applied either uniformly in both agricultural sectors or to only one of the two. This property permits examination of the differences in actual and potential productivity between irrigated and other agricultural areas.

Technical change alters product supply and factor demand directly in the sector to which the shock applies and indirectly in all sectors through adjustments in factor and product markets. The values of these changes are obtained by the requirement that changes in supply equal changes in demand in the markets for labor, capital, and the nontradable good. Because the model is oriented to the short run, changes in mobile factor supplies are constrained to zero; thus, any change in the functional distribution of income can be read directly from changes in factor prices.

The technical change specification in the model is adapted from Quizon and Binswanger (1983). In their analysis the rate and the bias of technical change in sector s can be obtained from the values of the factoral rates of technical change, A'_{is} , for each factor i . The A'_{is} terms are the percentage changes in factor demand due to technical progress, evaluated at constant prices and output:

$$A'_{is} = \frac{\partial X_{is}}{\partial t} \frac{1}{X_{is}},$$

where X_{is} denotes the demand for factor i in sector s , and t denotes time. The overall rate of technical change T'_s is the cost-share-weighted average of the A'_{is} terms:

$$T'_s = \sum_{i=1}^n \alpha_{is} A'_{is},$$

where α_{is} denotes the share of factor i in total costs in sector s . For ease of exposition, in the subsequent discussion the sector subscripts will be suppressed.

A'_i and T' are defined when output is held constant, as is the case in a cost function. Quizon and Binswanger (1983) derive the following expressions for E'_i and E'_y , the profit function counterparts of A'_i and T' above:

$$(1) \quad E'_i = \beta_{iy} T' + A'_z - A'_i; \text{ and}$$

$$(2) \quad E'_y = \beta_{yy} T' + A'_z,$$

where β_{ij} is the elasticity of quantity i with respect to the price of j , and A'_z is the factoral rate of technical change of land, the fixed factor in

the profit function. In this model the E'_i 's and E'_y in the agriculture sector supply and factor demand equations are replaced by the right-hand sides of equations (1) and (2). Selecting values of the A'_i 's and T' permits the simulation of any combination of technical change rates and biases, as will now be shown.

Factor-neutral technical progress in any sector is characterized by equality of each of the factoral rates of technical change A'_i and the overall rate of technical change T' . This equality implies the following changes in factor demand and output in that sector:

$$(3) \quad \begin{aligned} E'_l &= \beta_{ly} T' \geq 0; \\ E'_k &= \beta_{ky} T' \geq 0; \text{ and} \\ E'_y &= (\beta_{yy} + 1) T' \geq 0. \end{aligned}$$

Technical progress which substitutes capital for labor with no change in output or in the productivity of land ($A'_z = T' = 0$) yields

$$(4) \quad \begin{aligned} E'_l &= -A'_l \leq 0; \\ E'_k &= \left(\frac{\alpha_l}{\alpha_k} \right) A'_l \geq 0; \text{ and} \\ E'_y &= 0. \end{aligned}$$

These expressions show that the *ceteris paribus* effect of the capital-using bias of technical change is to reduce labor demand and raise capital demand. Reversal of k and l subscripts will yield expressions for the effects of a capital-saving, labor-using bias.

The third illustrative technical change involves the substitution of labor for land, with no change in the rate of technical progress or in its bias with respect to capital ($A'_k = T' = 0$). Because land is a specific factor and constant returns to scale have been imposed, at constant prices any increase in land productivity automatically ensures an increase in output even though the physical land area remains unchanged. In this case the technical change shifters are

$$(5) \quad \begin{aligned} E'_l &= \left(\frac{\alpha_z}{\alpha_l} + 1 \right) A'_z \geq 0; \\ E'_k &= A'_z \geq 0; \text{ and} \\ E'_y &= A'_z \geq 0. \end{aligned}$$

According to these definitions, the individual components of any technical change which both increases output and alters factor proportions can be isolated by applying shocks to individual

variables: T' for factor neutrality, A'_i for labor-capital bias, and A'_l for land-saving, labor-using bias. The nonneutral changes (4) and (5) impart only pure factor substitution effects; but, since the model is linear in percentage changes, either or both may be combined with the results for neutral change to simulate the effects of progress which both raises output and alters the proportions in which factors are demanded at constant prices. Any combination of technical progress and factor bias can be represented in the manner just described so long as the number of technical change variables (T' and the A'_i terms) in each sector is equal to the number of factors employed in that sector.

The effects of other combinations of factor bias and output increase could readily be specified; for the illustrative purpose served by this model, however, the three just described are adequate. In the experiments simulating unbalanced growth in agriculture, technical changes in sector 1 only are described, but an equivalent methodology applies to sector 2 because technical change mechanisms are algebraically identical in both sectors. The effects of a constant rate and bias of technical progress across both agricultural sectors are also presented and discussed below.

Household Income and Consumption

The model contains seven illustrative groups of households, distinguished both by asset ownership and by their consumption behavior. They are summarized in table 3. The small farmer categories (H6 and H7) include tenant farmers on fixed-rent leases because they also derive benefits from increased land productivity when output rises. The definitions of households used in this study have no direct equivalents in official Philippine statistics, in which households are grouped by levels of income and expenditure. Such information on consumption patterns as can be gleaned from official statistics (NEDA, various years) and independent studies (Pante; Lluch, Powell, and Williams; Kravis, Heston, and Summers) has been used to construct typical sets of consumer demand parameters for each of the household groups. Shares of total household income contributed by each factor owned by the k th household group are shown in table 4. These data have been constructed synthetically with the aim of providing a contrast between stereotypical groups.

The extent to which the incomes of farmers and landlords are derived from the earnings of

land reflects the assumption that they own or lease farm land in which they have made some long-term investment; for this reason their fortunes are closely tied to profitability in the sector in which their land belongs. The sector-specificity of farmers' incomes from land is not contradicted by their also deriving income from each of the mobile factors: their endowments of labor, for example, may be switched between on-farm and off-farm employment as changes in sectoral profitability dictate.

Laborers do not share in the ownership of fixed assets; their incomes are affected by a change in a particular sector only insofar as it affects the economy-wide demand for their services. Households owning specific factors outside agriculture, in services (H2) and manufacturing (H3), have asset ownership positions roughly equivalent to those of landlords: although they derive some income from their ownership of mobile capital, the greater portion comes from returns to specific factors. Changes in sectoral profitability are, thus, the main determinants of changes in these households' relative prosperity.

Solution Procedure

The model is implemented using the GEM-PACK software package (Codsí and Pearson).

Table 3. Household Categories and Characteristics

Code	Description	Asset Ownership
H1	Landless laborers	Mobile labor only
H2	Service sector capitalists	Specific factor in sector 3; some mobile capital
H3	Manufacturing sector capitalists	Specific factors in sector 4; some mobile capital
H4	Landlords in irrigated area (sector 1)	Most irrigated land, some mobile capital
H5	Landlords in unirrigated area (sector 2)	Most unirrigated land, some mobile capital
H6	Small farmers in irrigated area (sector 1)	Some irrigated land, some mobile labor and capital
H7	Small farmers in unirrig. area (sector 2)	Some unirrigated land, some mobile labor and capital

Note: Exact quantities of factors owned by each group are shown in table 4.

Table 4. Factor Ownership and Factor Shares in Household Income

Household Group ^a	δ_{hi}	δ_{hk}	γ_{h1}	γ_{h2}	γ_{h3}	γ_{h4}	$\Sigma_i \delta_{hi} + \Sigma_i \gamma_{hi}$
<i>H1</i>	1.0	0.0	0.0	0.0	0.0	0.0	1.0
(ϕ_{1i})	(0.4) ^b	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	
<i>H2</i>	0.0	0.3	0.0	0.0	0.7	0.0	1.0
(ϕ_{2i})	(0.0)	(0.2)	(0.0)	(0.0)	(1.0)	(0.0)	
<i>H3</i>	0.0	0.3	0.0	0.0	0.0	0.7	1.0
(ϕ_{3i})	(0.0)	(0.5)	(0.0)	(0.0)	(0.0)	(1.0)	
<i>H4</i>	0.0	0.5	0.5	0.0	0.0	0.0	1.0
(ϕ_{4i})	(0.0)	(0.1)	(0.7)	(0.0)	(0.0)	(0.0)	
<i>H5</i>	0.0	0.4	0.0	0.6	0.0	0.0	1.0
(ϕ_{5i})	(0.0)	(0.1)	(0.0)	(0.7)	(0.0)	(0.0)	
<i>H6</i>	0.3	0.4	0.3	0.0	0.0	0.0	1.0
(ϕ_{6i})	(0.3)	(0.05)	(0.3)	(0.0)	(0.0)	(0.0)	
<i>H7</i>	0.4	0.2	0.0	0.4	0.0	0.0	1.0
(ϕ_{7i})	(0.3)	(0.05)	(0.0)	(0.3)	(0.0)	(0.0)	
$\Sigma_h \phi_{hi}$	(1.0)	(1.0)	(1.0)	(1.0)	(1.0)	(1.0)	

Source: See text.

^a Household groups (defined in text) are *H1*, laborers; *H2*, owners of sector 3's specific factor; *H3*, owners of sector 4's specific factor; *H4*, landlords in sector 1; *H5*, landlords in sector 2; *H6*, farmers in sector 1; and *H7*, farmers in sector 2.

^b Figures in parentheses indicate ϕ_{hi} , the proportion of the economy's endowment of each factor *i* owned by household group *h*.

In table 1 the complete model is shown to consist of $(4s + 4h + 6)$ equations and $(5s + 4h + 17)$ variables. Because tradables' prices are assumed constant, their percentage changes are set to zero in the model's equations. Closure is achieved by specifying as exogenous the technical change shifters E'_i and E'_y and the endowments of capital, labor, and fixed factors \bar{K}' , \bar{L}' , and Z'_1 to Z'_4 . This leaves $(4s + 4h + 6) = 50$ endogenous variables and the same number of equations, permitting the model to be solved.

Sensitivity Analysis

Each of the parameters obtained by econometric analysis is estimated with error resulting from random variations in the data. Many other parameters of the model are derived not from econometric estimates but from independent information such as national accounts data. Considerable uncertainty over the values of these parameters is justified.

The possibility of errors in parameter values raises two issues. One is that the means by which some values have been obtained leaves open the possibility that they could be substantially different from their true values. The second is that the model may be extremely sensitive even to small changes in some parameter values. An empirical investigation of the second issue provides guidance as to the importance of the first. If the solution values of the model's dependent

variables are very insensitive to changes in parameter values, then uncertainty about the exact value of a parameter need not greatly reduce confidence in the outcomes of simulation experiments.

Sensitivity elasticity is defined as the percentage change in a solution value given a 1% change in the value of a parameter or parameter group (Pagan and Shannon, Anderson and Warr). Sensitivity elasticities were computed for each variable with respect to a 10% change in each parameter for a neutral technical change shock in sector 1. With 43 variables and 197 parameters in the model, 8,471 such elasticities could in principle have been calculated. The actual number computed was somewhat smaller (5,461) because it excluded changes in elasticities relating to changes in tradables' prices—these were held constant in our model—and parameters whose initial values were zero.

The solutions obtained for the base (neutral technical progress) simulation run are relatively insensitive to large changes in both estimated and guesstimated parameter values. Of the 5,461 sensitivity elasticities computed, only 93 (1.7%) exceeded 1.0 in absolute value; among these, the modal value was less than 2.0. These results indicate that the model is relatively robust with respect to its parameters.³

³ More detail on the procedure and results for sensitivity analysis is available from the authors on request.

Experimental Results

In this section we report and discuss the results of technical progress shocks conducted under a variety of assumptions concerning sectoral access to new technologies and factor biases. The purpose of these experiments is to expose the mechanisms by which functional and household income distributions are altered by a shock. First, counterfactual factor-neutral shocks are decomposed into their economic components. Second, the impacts of factor biases in technical change are studied in isolation from one another using hypothetical values of the technical change parameters T' , A'_i and A'_j . Although the data used in the model are based on Philippine data, the underlying model is, by design, highly stylized.⁴

The results are presented below by decomposing the percentage change in each endogenous variable into two components: the resource movement effect and the spending effect. The resource movement effect describes changes due to adjustments other than the spending of additional income generated by technical progress. Since the spending effect operates through demand for the non-tradable good (Corden and Neary; Cassing and Warr) the resource movement effect is calculated in these simulations by setting the expenditure elasticity of demand for the non-tradable, η_3 , equal to zero, which nullifies the spending effect. The resource movement effect is further divided into the factor market effect and the real appreciation effect.⁵ The former is calculated by holding the *numéraire* price of the nontradable constant, which nullifies the real appreciation effect. The log-linearity of the model ensures that the effects of an exogenous shock can always be represented as the sum of these components. In this manner the analysis exposes the relative importance of factor market and consumption linkages, and of the agriculture/nonagriculture terms of trade, in determining the distributional outcomes of technical change shocks. By revealing the major adjustment mechanisms, decomposed results reduce substantially the "black box" nature of CGE experiments, thus helping nontechnical users of such

models to gain an intuitive understanding of the results obtained.

The absolute prices of the two internationally traded commodities, the agricultural good and the manufactured good, are held constant throughout the analysis. Other prices, all of which may change relative to those of these *numéraire* goods, are measured in terms of the *numéraire* goods, and are thus called *numéraire* prices.

Balanced Technical Progress

An intuitive grasp of the results is made easiest by first considering the case in which a uniform, factor-neutral technical change occurs in sectors 1 and 2 together. Percentage changes in selected endogenous variables resulting from a uniform 10% increase in the productivity of all factors in agriculture are reported in columns (1) to (5) of table 5. For prices, the real changes reported are obtained by dividing changes in *numéraire* prices by the GDP-share-weighted sum of changes in commodity prices. These will be called changes in deflated prices.

In the balanced technical change case, sectors 1 and 2 have equal access to new technology. Output and factor demands rise in both sectors as the result of technical progress. Because producers in sector 1 (irrigated agriculture) are more price responsive than their counterparts in sector 2 (unirrigated agriculture) (Coxhead), output and associated factor demands rise by somewhat more in the former than in the latter. At constant output prices (column 1), the additional factor demand in agriculture is met by drawing capital and labor out of services and manufacturing. Real prices of the mobile factors rise accordingly, with the greater increase accruing to labor, because that is the mobile factor used relatively intensively in agriculture. Real returns to fixed factors in the expanding agricultural sectors rise by more than the amount of the technical change (as predicted for lower-dimensional models by Jones), while returns to fixed factors in other sectors fall as increases in mobile factor prices reduce profitability. At constant output prices, the technical change shock raises the incomes of the owners of factors used relatively intensively in agriculture, labor and land, with the greatest share of the increases captured by owners of the factors in least elastic supply. Thus, among the households, the greatest real gains are won by landlords, with lesser gains to farmers (who own some land, labor, and capital),

⁴ A 42-sector general equilibrium model of the Philippine economy is presently under construction by the authors.

⁵ A real appreciation is a rise in the price of nontradables relative to the prices of tradable commodities, i.e., a rise in the real exchange rate. The real appreciation effect as defined here refers to the *ceteris paribus* effect of a rise in the relative price of nontradables, net of the effects of factor market adjustments and income effects.

Table 5. Effects of Neutral Technical Change Shocks in Agriculture (%)

Endogenous Variable	Balanced Growth ($T'_1 = T'_2 = 10.0$)					Unbalanced Growth ($T'_1 = 10.0$) Total Change (6)
	Resource Movement Effect			Spending Effect (4)	Total Change (5) = (3) + (4)	
	Factor Markets Effect (1)	Real Apprec'n Effect (2)	Total Resource Movement Effect (3) = (1) + (2)			
Labor demand						
Agriculture 1	3.53	-0.28	3.25	-3.65	-0.40	1.85
Agriculture 2	0.76	-0.07	0.69	-1.00	-0.31	-0.68
Services	-1.32	0.49	-0.83	6.36	5.53	2.74
Manufacturing	-1.88	-0.36	-2.24	-5.31	-7.55	-3.93
Capital demand						
Agriculture 1	4.55	-0.33	4.22	-4.18	0.04	2.24
Agriculture 2	0.58	-0.03	0.55	-0.38	0.17	-0.20
Services	-0.69	1.05	0.36	13.51	13.87	6.82
Manufacturing	-0.49	-0.66	-1.15	-8.59	-9.74	-5.25
Product supply						
Agriculture 1	14.86	-0.28	14.58	-3.49	11.09	13.22
Agriculture 2	11.60	-0.07	11.53	-0.90	10.63	-0.61
Services	-1.47	1.02	-0.45	13.20	12.75	6.29
Manufacturing	-1.47	-0.64	-2.11	-8.22	-10.33	-5.47
Deflated prices						
Labor	1.86	0.03	1.89	0.35	2.24	1.33
Capital	1.08	0.00	1.08	0.03	1.11	0.80
Services	0.00	1.04	1.04	13.39	14.43	7.34
Fixed factor 1	22.69	-1.58	21.11	-20.41	0.70	12.31
Fixed factor 2	18.29	-1.27	17.02	-16.35	0.67	-10.05
Fixed factor 3	-3.57	3.43	-0.14	44.13	43.99	21.94
Fixed factor 4	-3.30	-2.12	-5.42	-27.16	-32.58	-17.10
Household incomes						
H1	1.86	0.41	2.27	5.23	7.50	4.01
H2	-2.17	2.53	0.36	32.59	32.95	16.52
H3	-1.99	-1.34	-3.33	-17.32	-20.65	-10.80
H4	11.89	-0.66	11.23	-8.50	2.73	7.48
H5	11.41	-0.63	10.78	-8.11	2.67	-4.78
H6	7.80	-0.01	7.79	-0.07	7.72	7.67
H7	8.28	-0.04	8.24	-0.46	7.78	-0.07

Note: For full details of simulation results see Coxhead and Warr (1989).

and considerably lower gains accruing to laborers.

The second component of the resource movement effect is the real appreciation effect, shown in column 2 of the table. Relative to all commodity prices, the price of the nontradable rises by a little more than 1%. This rise in profitability in the nontradable sector draws mobile factors away from other sectors: the demand for capital and labor rises in sector 3 and falls in all other sectors. Output in the nontradables sector recovers about half of the ground lost from factor market adjustments: having fallen by 1.47%, the rise in P_3 reduces the decline in output to -0.45%. The deflated return to the specific factor in sector 3, which was cut by 3.57% by the factor market adjustments, rises as the result of real appreciation by 3.43%: the net impact of

the resource movement effect on the return to this factor is thus virtually zero. This result stands in contrast to that for the manufacturing sector. De-industrialization takes place both directly through the movement of factors into agriculture and indirectly through the real appreciation. Output in sector 4 falls by 2.11% as the result of the resource movement effect of the shock.

In real terms, the rise in nontradables' price resulting from the resource movement effect has no effect on wages or returns to mobile capital but raises the return to the factor specific to the nontradables sector. As the zero pure profit conditions require, the change in P_3 is "trapped between," i.e., is a factor share-weighted average of, changes in the prices of factors used in the production of nontradables.

Real changes in household income are ob-

tained by deflating *numéraire* incomes by household-specific commodity price indices. On its own, the real appreciation component of the resource movement effect redistributes some of the incremental income derived from the technical progress away from agriculturalists towards the owners of the factor specific to nontradables: the real incomes of households in the group *H3* rise. Laborers also gain. Rises in returns to their endowments of mobile capital are insufficient to offset the losses experienced by landlords and the owners of the specific factor in manufacturing.

The spending effect boosts output and factor demand in nontradables primarily at the expense of the manufacturing sector, although agricultural factor demands and output levels also fall slightly. Increased demand raises the real price of nontradables by 13.4%, considerably more than the rises in returns to mobile factors and more than ten times greater than the real appreciation associated with the resource movement effect. This contributes to a final outcome for deflated factor returns, which confirms a rise in wages (2.24%) about double that in returns to mobile capital (1.11%). A large rise in the return to the factor specific to nontradables from the spending effect yields an overall rise in that factor's real price of 44%. Real returns to manufacturing's specific factor fall by 33%. Of greatest interest is the impact of the spending effect on real returns to agricultural land: the gains from the technical change shock are virtually cancelled by subsequent changes in mobile factor prices and in the price of nontradables.

Unbalanced Technical Progress

In the developing countries of Asia, a disproportionately large part of public investment in agricultural research and extension has been directed toward the improvement of agricultural productivity in irrigated areas. The rate of technical progress in these areas has consequently been much greater than that in areas without access to irrigation (David and Otsuka). The changes in prices and incomes resulting from a technical change shock restricted to one agricultural region (sector 1) are shown in column 6 of table 5. Most of the additional income generated directly by the technical progress is captured by landowners in sector 1, especially by landlords. As in the balanced growth case, the spending effect substantially increases profitability in the nontradable sector. In sector 2, ris-

ing factor prices reduce profitability, driving down output and the rents paid to sector-specific land. The real appreciation, the indirect component of the resource movement effect, appears to all agricultural producers as a fall in the price of their output relative to the general commodity price level. Without the opportunities for unit cost reductions provided by the new technologies, this further reduces profits and output in the lagging agricultural sector. Moreover, the operation of the spending effect on the demand for nontradables spurs further factor price rises and another reduction in the relative price of farm produce, and so further reduces profits in sector 2. These losses inflicted on the lagging agricultural sector by its lack of access to new technologies are not captured by most empirical partial equilibrium analyses.

Unbalanced Growth and the Distribution of Income

The functional income distribution. The changes in real factor returns shown in table 5 enable us directly to compare changes in real income between the owners of sector-specific land endowments in the two agricultural sectors. A second comparison of interest in developing countries is between the income of labor and the aggregate return to land. Because labor is mobile between sectors, it is appropriate to compare the change in its income with the change in income accruing to all agricultural land, rather than to land in a single sector. Changes in the aggregate return to land are calculated as the weighted average of changes in land prices in sectors 1 and 2, calculated from columns 5 and 6 of table 5. The weight used is θ , the value of production in irrigated areas as a proportion of the total value of agricultural production. In 1980, the reference period of the Philippine data used in simulations, the value of θ was 0.2087.

When both agricultural sectors grow, the difference between sectors 1 and 2 in the changes in real returns to land is negligible. The θ -weighted average of these changes, the percentage change in the aggregate return to land, is calculated as

$$z'_{agg} = \theta(0.70) + (1 - \theta)(0.67) = 0.68.$$

Under unbalanced growth the gap is naturally much wider, composed not only of an increase in the real return to land in sector 1 but also of a fall in the real return to land in sector 2. Consequently, if growth is unbalanced the change

in aggregate returns to land is actually negative:

$$z'_{agg} = \theta(12.31) + (1 - \theta)(-10.05) = -5.38.$$

These results are summarized in column 1, rows 1 and 2, of table 6. The second column of the table shows that when the rate of technical progress is uniform in both sectors, owners of land lose slightly relative to laborers. At 1.56%, the difference between changes in land and labor returns is small enough for the effect of balanced growth to be described as neutral with respect to the labor-land distribution of income. When technical progress occurs only in sector 1, however, labor gains relative to landowners: subtracting the aggregate landowners' loss from the 1.33% rise in labor income yields a margin of 6.71%.

The mobility of labor, combined with the fact that labor gains from both the factor and product price components of the resource movement effect as well as from the spending effect, ensures that among factors employed in agriculture, labor captures a substantial share of the total gains from the factor-neutral technical progress. Whether or not labor gains relative to capital depends on the relative factor intensities of the sector experiencing technical progress and of the sector producing nontradables, and on the rate at which each sector grows. Under balanced growth in the labor-intensive agricultural sectors, the increase in labor demand is much greater than that of capital; wages thus rise relative to capital's price. When only sector 1 experiences technical progress the relative wage rise is halved, to 0.53%.

Analyses of technical progress which consider only the distribution of gains among factors employed wholly within agriculture ignore the impact of the spending effect on factor price relativities. That this omission can result in completely different interpretations of the dis-

tribution of factor incomes following technical progress is evident from rows 3 and 4 of table 6, which show changes in relative factor prices resulting only from the resource movement, or direct effect, of the technical change shock. Column 2 of the table shows that on its own, the resource movement effect reduces wages relative to land rents by 16% if growth affects both agricultural sectors and by 2% if it is restricted to sector 1. In Quizon and Binswanger's (1986) model of agricultural growth in India, nonagricultural production and technology are by assumption exogenous, which implicitly requires that agricultural growth has no effect through backward or forward linkages on factor demand in other sectors. Their finding that technological improvements in agriculture do less to improve the incomes of India's rural poor than do increases in nonagricultural employment (Quizon and Binswanger 1986, p. 105) must be evaluated in the light of their assumption that labor markets are dualistic in structure.⁶

The household income distribution. The household distribution of income is found by mapping the functional distribution onto the pattern of asset ownership by which household groups are defined. Under balanced agricultural growth, landlords in sectors 1 and 2 experience similar rises in real income, as do the two small farmer groups (table 5, column 5). In the unbalanced growth case, sector 2's landlords and farmers lose in real income because of the fall in returns to their land (column 6). As with factor returns, it is relevant to contrast the house-

⁶ Among the strongest backward and forward linkages of agriculture in developing countries are those to rural-based, labor-intensive activities with a high nontradables component. These include grain milling, food processing, transport and storage, input supply and small-scale implement manufacture and supply.

Table 6. Comparative Changes in Factor Returns from Neutral Technical Progress (%)

	Land Price Ratio ($z'_1 - z'_2$) (1)	Wages Relative to Aggregate Land Price ($w' - z'_{agg}$) (2)	Wages Relative to Capital Price ($w' - r'$) (3)
Total effect of technical change shock			
Balanced growth	0.03	1.56	1.13
Unbalanced growth	22.34	6.71	0.53
Resource movement effect only of technical change shock			
Balanced growth	4.09	-15.98	0.81
Unbalanced growth	24.57	-2.04	0.38

hold income changes experienced by laborers against those of landlords and farmers in sectors 1 and 2 together. The change in the real income of equivalent groups in both agricultural sectors is computed as the θ -weighted sum of changes in real income of each household group. These comparisons are shown in table 7.

In the balanced technical progress experiment real returns to land, labor, and capital all rise.⁷ Table 7 shows that farmer households, whose endowment bundles include some of each factor, gain income relative to laborers and also relative to landlords, who own land and mobile capital. Landlords lose income relative to laborers. When technical progress is restricted to sector 1, however, the gains of laborers relative to both farmers and landlords are positive, even though they lose relative to landlords and farmers in sector 1, because of the losses incurred by landlords and farmers in sector 2.

Income distribution between households owning factors employed in agriculture appears quite

stable when rates of (neutral) technical progress are equal in both agricultural sectors. The relative gains of landowners from the resource movement effect are offset by gains accruing to owners of mobile factors when the spending effect is taken into account. Under unbalanced growth, disparities in income changes within each sector are greater; for example, laborers lose substantially relative to landlords in sector 1. This is so primarily because the spending effect, and consequent increases in labor demand and reductions in the real price of agricultural output, is much smaller relative to the balanced growth case.

These distributional results illustrate an important feature of the general equilibrium approach to the distribution of gains from technical progress. Partial equilibrium models and some "limited" general equilibrium models (Quizon and Binswanger 1986) attempt to capture the effects of changes in relative prices by assuming that the demand for agricultural output is price inelastic. This approach fails to take account of the effects on nonagricultural production and factor demand resulting from changes in the prices of intersectorally mobile factors. The true general equilibrium impact of technical change is thus captured only partially through commodity prices. The effects of changes in mobile factor prices are also important for the results in agriculture. By contrast with partial equilibrium approaches, in a properly specified general equilibrium model demand for the good produced in the sector experiencing technical progress need not be inelastic in order for the greater part of the gains from growth to be captured by groups other than owners of factors specific to that sector. It is sufficient that the mobile factors used in the expanding sector are also used to produce another commodity (in this

⁷ In table 5, column (5), the real prices of both mobile factors are seen to rise following a factor-neutral technical change shock. This harmony of the interests of labor and capital contrasts with the conflict implicit in models with fewer specific factors (see Cassing, and Warr, and Corden and Neary). We can summarize the preconditions for such a result. If factor supply elasticities are equal, then for the change in the wage-rental ratio $w'-r'$ to be unambiguously positive (negative) requires that both the sector experiencing technical progress and the sector producing nontradables, which grows as a result of the spending effect, be labor-intensive (capital-intensive) in terms of the two mobile factors. In this simulation model sector 1, the so-called "booming" sector, is labor intensive relative to capital. The resource movement effect of the technical progress thus raises $w'-r'$. Sector 3, however, is capital intensive relative to labor, so the spending effect lowers $w'-r'$. When the general equilibrium effects of the technical progress "boom" are taken into account, the interests of owners of the mobile factors are not opposed. Rather, their combined interests are in opposition to those of the owners of specific factors because any rise in specific factor endowments raises w and r relative to all specific factor prices and vice versa.

Table 7. Comparative Changes in Real Household Incomes from Neutral Technical Progress (%)

Income Gains of:	Relative to:	Balanced Growth	Asymmetric Growth
Laborers (H1)	Landlords 1 (H4)	4.77	-3.47
Laborers	Landlords 2 (H5)	4.83	8.79
Laborers	All landlords	4.82	6.23
Laborers	Farmers 1 (H6)	-0.22	-3.66
Laborers	Farmers 2 (H7)	-0.28	4.08
Laborers	All farmers	-0.27	2.47
Farmers 1	Landlords 1	4.99	0.19
Farmers 2	Landlords 2	5.11	4.71
All farmers	All landlords	5.09	3.76

case, nontradables) whose price is endogenous and responsive to changes in domestic demand. The inflation of this price caused by income effects arising from the technical progress and resulting factor market adjustments is sufficient to reduce the real price of the good produced in the expanding sector.

Biased Technical Change

Equations (4) describe the constant-price effects of a technical change which substitutes capital for labor with no increase in output.⁸ Column 2 of table 8 summarizes the total change in prices and incomes from both the resource movement and spending effects of such a shock. The substitution of capital for labor in sector 1 results in a real wage decline of 1.67%, in spite of the substitution of labor for capital in all other sectors. Sectors facing elastic demand for their output (1, 2, and 4) benefit from falling (deflated) wages, and their outputs rise slightly.⁹ Households owning only labor lose from the labor-saving change: their real incomes fall by 1.49%. Farmers in sector 2 also derive much of their income from labor, and, because they do not

benefit as landowners from technical progress in sector 1, their real incomes fall slightly. Households not dependent on labor experience small rises in real income. For sector 1 farmers, the gains from rises in land returns just offset losses in their labor incomes. Because laborers and small farmers comprise most of the low-income families in developing countries, a labor-saving, capital-using technical change bias must lead to a worsening income distribution.

In the second nonneutral component of technical progress considered in table 8, labor is substituted for land. Declining per capita availability of arable land in the Philippines has spurred the adoption of a range of technologies which economize on its use. Irrigation and high-yielding grain varieties rely on substantial increases in the per hectare inputs of both labor and intermediate goods. Column 3 of table 8 reports the total changes in prices and incomes following the labor-using, land-saving shock described by equations (5). Because both mobile factors are complementary with land and the effective land area has been increased by the shock, real wages as well as returns to mobile capital increase. Because the change is labor using, wages increase substantially relative to capital's price (the difference is 2.44%), and real returns to fixed factors are greatly reduced. Together, these results ensure a strongly positive effect on income distribution, given that laborers and small farmers are the poorest household groups in the economy. Laborers' real incomes rise by 5.38%, and those of small farmers in the lagging sector rise by 0.69%.

⁸ Recent Philippine examples of this type of change include the mechanization of land preparation (Sison, Herdt, and Duff; McCoy) and of rice threshing (Hayami and Kikuchi); and direct seeding, which is the substitution of herbicides for transplanting and weeding labor when the rice crop is established by broadcasting seed instead of transplanting (Moody and Cordova).

⁹ Full details of changes in factor demand and commodity supply are reported in Coxhead and Warr.

Table 8. Effects of Neutral and Nonneutral Technical Change Shocks in Sector 1

Endogenous Variable	Factor-Neutral ($T'_1 = 10.0$) (1)	Labor-Saving, Capital-Using ($A'_1 = 10.0$) (2)	Labor-Using, Land-Saving ($A'_2 = 10.0$) (3)
Prices			
Labor	1.33	-1.67	5.08
Capital	0.80	0.22	2.64
Services	7.34	0.50	0.83
Fixed factor 1	12.31	0.57	-7.34
Fixed factor 2	-10.05	0.69	-5.59
Fixed factor 3	21.94	3.68	-6.64
Fixed factor 4	-17.10	0.38	-10.25
Household incomes			
H1	4.01	-1.49	5.38
H2	16.52	2.70	-3.75
H3	-10.80	0.39	-6.28
H4	7.48	0.46	-2.25
H5	-4.78	0.57	-2.19
H6	7.67	-0.02	0.74
H7	-0.07	-0.13	0.69

Some of the qualitative effects of technical change are predicted by single-sector or partial equilibrium models. For example, virtually any such model will predict that land-saving, labor-using technical change in agriculture will produce a redistribution of income from landlords to laborers (e.g., Ahammed and Herdt). Other important effects, in particular those involving intersectoral linkages, will be missed by these models. A good example is the dramatic fall in returns to land in sector 2 resulting from the effect on factor prices of technical change in sector 1.

Empirically Estimated Technical Change Shocks

Values of technical change parameters appearing in the above model have been estimated from Philippine agricultural data (Coxhead). The estimates reveal a strong land-saving, fertilizer-using bias and fairly rapid overall rate of technical change in irrigated environments, and a much slower rate of technical change with correspondingly smaller biases in unirrigated environments. These values (shown in table 9, scaled to simulate an exogenous 10% technical change shock) were employed in additional experiments. The results, summarized in table 10, bear a close resemblance to the hypothetical effects of a neutral shock restricted to sector 1, as shown in column (6) of table 5. The spending effect, and the real appreciation component of the resource movement effect are less important in determining distributional outcomes than are the factor market adjustments associated with the shocks. Because the overall rate of technical progress in sector 2 is so low, real returns to owners of sector 2 land fall dramatically. Real mobile factor prices rise, with w' exceeding r' by 1.8%. The magnitude of the land-saving bias ensures that the real incomes even of sector 1

landlords fall relative to laborers' incomes. Laborers' incomes rise by 11.14% relative to aggregate landlords' incomes, and by 5.7% relative to all farmers' incomes.

Considered separately from contemporary changes in factor endowments and from changes in relative prices arising from other sources, the direction of technical change in Philippine agriculture appears to have brought relatively greater benefits to poorer groups (farmers and laborers), and thus to have exerted a positive influence on income distribution. This is so in spite of a clear worsening in the profitability of production in sector 2; the losses in that sector are absorbed by households owning sector 2-specific land in proportion to their ownership of it. Sector 2's landlords lose heavily, and the incomes of farmers in poor agricultural areas fall relative to those of laborers and of farmers in well-irrigated areas.

Conclusion

This paper has explored the short-run income distributional effects of technical change in agriculture with the aid of a small, stylized general equilibrium model. The model distinguishes between commodities which are traded internationally and those which are not, between fixed and mobile factors of production, and between technical changes of different factor biases. An important feature of the model is that it also distinguishes between groups of agricultural producers with and without access to resources complementary with the new technology, notably irrigation.

The simulations of the general equilibrium effects of technical change draw attention to the fact that the interests of owners of fixed factors do not necessarily coincide; returns to different fixed factors, or to the same fixed factor used in different sectors, may, and frequently do, move in different directions. It is well known that in

Table 9. Values of Empirical Technical Change Shocks

Sector	Land A'_L	Labor A'_L	Fertilizer A'_K	Overall Rate T'
Agriculture 1 (irrigated)	14.826	6.265	-1.911	7.645
Agriculture 2 (nonirrigated)	1.031	-1.164	-0.019	0.267

Source: Computed from estimates in Coxhead (1989).

Note: For factors $i = Z, L, K$, $A'_i > 0$ implies that at constant prices and output, technical progress has increased the productivity of (i.e., has "saved") factor i .

Table 10. Effects of a Technical Change Shock with Empirically Estimated Technology Response Parameters

Endogenous Variable	Resource Movement Effect			Spending Effect (4)	Total Change (5) = (3) + (4)
	Factor Markets (1)	Real Apprec'n (2)	Total RM Effect (3) = (1) + (2)		
Prices					
Labor	4.99	0.08	5.07	0.10	5.17
Capital	3.32	0.01	3.33	0.01	3.34
Services	0.00	2.94	2.94	3.95	6.89
Fixed factor 1	12.67	-4.48	8.19	-6.03	2.16
Fixed factor 2	-4.12	-3.59	-7.71	-4.84	-12.55
Fixed factor 3	-9.98	9.69	-0.29	13.04	12.75
Fixed factor 4	-9.42	-5.96	-15.38	-8.03	-23.41
Household incomes					
H1	4.99	1.15	6.14	1.55	7.69
H2	-5.99	7.15	1.16	9.64	10.80
H3	-5.60	-3.80	-9.40	-5.11	-14.51
H4	7.99	-1.86	6.13	-2.51	3.62
H5	-1.15	-1.78	-2.93	-2.39	-5.32
H6	6.63	-0.02	6.61	-0.02	6.59
H7	1.01	-0.10	0.91	-0.14	-0.77

small open economies the interests of farmers are hurt by increases in manufacturing protection, so to find that technical progress in agriculture may injure producers elsewhere is not surprising. In the case of landowners in agricultural sectors not experiencing technical change, their losses as a result of technical progress elsewhere will prompt efforts to adopt similar technological innovations for their own farms, as has been pointed out by Quizon and Binswanger (1983). In practice, however, their gains from adoption are constrained by the lack of complementary inputs such as irrigation. In the Philippines, the adoption rate of modern rice varieties in rainfed lowland and upland areas is the highest in Asia. In spite of this, rainfed rice yields have grown only very slowly.

The results of this study point to three critical issues in determining the income distributional effects of technical progress. These are the price-setting mechanisms for intersectorally mobile factors, heterogeneity in the quality of sector-specific factors (in this paper, land), and the factor biases of technical progress. These three matters are precisely those in which the desire for simplicity in modeling and the weakness of the empirical base often lead researchers to make default assumptions. Factor prices are assumed either to be fixed, or to be determined wholly within the sector under study. Agricultural land is taken to be of homogenous quality, and technical change is taken to be neutral with respect to factors. The policy relevance of analyses of

the welfare and distributional impact of technical progress or equivalent changes would be enhanced by increased attention to these important issues.

[Received September 1989; final revision received August 1990.]

References

- Ahammed, Chowdhury S., and Robert W. Herdt. "Farm Mechanization in a Semi-Closed Input-Output Model: The Philippines." *Amer. J. Agr. Econ.* 65(1983):516-25.
- Anderson, Kym, and Peter G. Warr. "General Equilibrium Effects of Agricultural Price Distortions: A Simple Model for Korea." *Food Res. Inst. Stud.* 20(1987):246-63.
- Bautista, Romeo M. "Effects of Increasing Agricultural Productivity in a Multisectoral Model for the Philippines." *Agr. Econ.* 1(1986):67-68.
- Cassing, James H., and Peter G. Warr. "The Distributional Impact of a Resources Boom." *J. Int. Econ.* 18(1985):301-20.
- Codsi, George, and Ken Pearson. "An Overview of GEM-PACK—A Software System for Implementing and Solving Economic Models." IMPACT Project, GEM-PACK Document No. 22, University of Melbourne, 1988.
- Corden, W. Max, and J. Peter Neary. "Booming Sector and Deindustrialization in a Small Open Economy." *Econ. Rec.* 92(1982):825-48.
- Coxhead, Ian A. "Technical Change, Irrigation and Factor

- Demand Elasticities in Philippine Agriculture." *Working Papers in Trade and Development*, No. 89/5. Canberra: Australian National University, 1989.
- Coxhead, Ian A., and Peter G. Warr. "Technical Change in Agriculture and the Distribution of Income: A General Equilibrium Model for the Philippines." *Working Papers in Trade and Development* No. 89/7. Canberra: Australian National University, 1989.
- David, Cristina C., and K. Otsuka. "Differential Impact of Technical Change Between Favorable and Unfavorable Areas." Paper presented at First Workshop on the Differential Impact of Modern Rice Technology on Favorable and Unfavorable Production Environments, International Rice Research Institute, Los Banos, Philippines, March 1987.
- Evenson, Robert E. "Infrastructure, Output Supply and Input Demand in Philippine Agriculture: Provisional Estimates." *J. Philippine Develop.* 23(1986):62-76.
- Hayami, Yujiro, and Robert W. Herdt. "Market Price Effects of Technological Change on Income Distribution in Subsistence Agriculture." *Amer. J. Agr. Econ.* 59(1977):245-56.
- Hayami, Yujiro, and Masao Kikuchi. *Asian Village Economy at the Crossroads: An Economic Approach to Institutional Change*. Tokyo: University of Tokyo Press, 1981.
- Johansen, L. *A Multi-Sectoral Study of Economic Growth*. Amsterdam: North-Holland Publishing Co., 1960.
- Jones, Ronald W. "A Three-Factor Model in Theory, Trade and History." *Trade, Balance of Payments and Growth*, ed. J. Baghwati et al., pp. 3-21. Amsterdam: North-Holland Publishing Co., 1971.
- Jones, Ronald W., and Stephen J. Easton. "Factor Intensities and Factor Substitution in General Equilibrium." *J. Int. Econ.* 15(1984):65-99.
- Kravis, Irving B., Alan W. Heston, and Robert Summers. "The Share of Services in Economic Growth." *Global Econometrics: Essays in Honour of Lawrence B. Klein*. ed. F. G. Adams and B. G. Hickman, pp. 188-218. Cambridge MA: MIT Press, 1983.
- Lipton, Michael, and Richard Longhust. *New Seeds and Poor People*. London: Unwin-Hyman, 1989.
- Lluch, Constantino, Alan A. Powell, and Ross A. Williams. *Patterns in Household Demand and Saving*. New York: Oxford University Press for the World Bank, 1977.
- McCoy, Alfred F. "An Agricultural Revolution: The Marcos Regime's Impact on Negros and the Sugar Districts." *The Philippines After Marcos*, ed. R. J. May and F. Nemenzo. London and Sydney: Croom Helm, 1988.
- Mendoza, Maria N., James A. Roumasset, and Ramon L. Clarete. "Planning Policy with General Equilibrium Models: The Effect of Tariffs in the Philippines." Resource Systems Institute Work. Pap. No. WP-83-19, East-West Center, University of Hawaii, 1983.
- Moody, Keith, and Violeta G. Cordova. "Wet-Seeded Rice." *Women in Rice Farming*, pp. 356-78. Los Banos, Philippines: International Rice Research Institute, 1985.
- National Economic and Development Authority (NEDA). *Philippine Statistical Yearbook*. Quezon City, various years.
- Pagan, A. R., and J. H. Shannon. "How Reliable Are ORANI Conclusions?" Centre for Economic Policy Research Disc. Pap. No. 130, Australian National University, 1985.
- Pante, Filogo. "Consumer Demand Functions: An Empirical Evaluation of Alternative Functional Forms." *Philippine Econ. J.* 18(1979):95-120.
- Quizon, Jaime B., and Hans P. Binswanger. "Income Distribution in Agriculture: A Unified Approach." *Amer. J. Agr. Econ.* 65(1983):526-38.
- . "Modeling the Impact of Agricultural Growth and Government Policy on Income Distribution in India." *World Bank Econ. Rev.* 1(1986):103-48.
- Sison, Jerome F., Robert W. Herdt, and Bart Duff. "The Effects of Small Farm Mechanization on Employment and Output in Selected Rice-Growing Areas in Nueva Ecija, Philippines" *J. Philippine Develop.* 12(1985):29-83.

Modeling Agricultural Growth Multipliers

Steven Haggblade, Jeffrey Hammer, and Peter Hazell

Agriculture's potential as an engine of third world growth depends, in large part, on the size of the production and consumption linkages it stimulates in rural regions. Current estimates of agricultural growth multipliers within rural regions vary widely, not only because economic structures differ across regions but also because the array of fixed-price models most commonly used embody widely differing basic assumptions. Among fixed-price models, the semi-input-output formulation projects the most plausible multipliers. But even they overstate the magnitude of growth multipliers by 10% to 25% according to the price-endogenous model developed here.

Key words: agricultural development, growth, multipliers.

Studies of the indirect consequences of the green revolution place regional agricultural income multipliers anywhere between 1.3 and 4.3. That is, a one dollar increase in technologically induced agricultural income generates an additional \$0.30 to \$3.30 in other sectors of a rural region (Ahmed and Herdt; Bell, Hazell, and Slade; Haggblade, Hazell, and Brown; Krishna; Little and Doeksen; Mirakhor and Orazem; Haggblade and Hazell). Although variations in economic structure across regions explain part of this wide range, differences in models and their assumptions may account for a greater share of the discrepancy.

And the magnitudes matter. Debate over agriculture's potential as an engine of third world growth hinges on the magnitude of the production and consumption linkages it generates (Hirschmann, Johnston and Kilby, Mellor). The linkages also play a central role in evaluating the equity implications of agriculture-led growth (Mellor and Johnston). In addition to their considerable dispersion, existing multiplier estimates may be overoptimistic. The models used to date—input-output, economic base and semi-input-output models—compute Keynesian demand-driven multipliers. All assume a perfectly elastic supply of nontradables. But if nontradable supply is not perfectly elastic, shifts in de-

mand will lead to higher prices and a reduction in estimated quantity increases (fig. 1). A price-endogenous model is required to capture these price effects and the resulting dampening of growth multipliers.

This paper aims to help quantify more precisely the magnitude of agricultural growth multipliers. It does so in two ways. First, the paper reviews the array of fixed-price models presently in use. It explores how they are related and examines analytically why they produce different multiplier estimates, even for identical economies. Second, it introduces a price-endogenous model that relaxes two key simplifying assumptions of the current fixed-price models: (a) it accommodates an upward-sloping supply curve for nontradables, and (b) it allows for substitution among inputs rather than insisting on fixed-coefficient Leontief production technology. The price endogenous model is then used to estimate the dampening effect of an upward-sloping supply of nontradables and to examine the consequences of input substitution for agricultural growth multipliers. To facilitate comparison, the paper applies the full complement of fixed-price and price-endogenous models to three specific rural regions to illustrate how differences in model assumptions affect multiplier estimates.

Steven Haggblade is an economist attached to Bojida Associates and is based in Dhaka, Bangladesh. Jeffrey Hammer and Peter Hazell are economists with the World Bank.

The views, findings, interpretations, and conclusions of this paper are the authors' and should not be attributed to the World Bank.

The authors are grateful to Mary Lovely and David Greytak for helpful discussions.

Fixed-Price Models

Most multipliers to date have been estimated using fixed-price models: economic base (EB), in-

put-output (I-O) and semi-input-output (S-I-O) models. To compute these multipliers, fixed-price models classify economic activities as tradable (T) or nontradable (N) according to the region's trade possibilities with the larger national economy and the rest of the world. For rural regions in developing countries, agricultural food grains or cash crops usually constitute the major tradable commodity, while nontradables usually include perishable agricultural goods such as fruits and meats, rural nonfarm services, and many rural manufactured goods.

All fixed-price models share three common assumptions. First, they view increased production of tradable goods as the driving force in regional economic growth. Second, they each model production as fixed-coefficient Leontief technology. And, third, they assume constant prices for both tradable and nontradable output.

To justify the constant output price for nontradables, practitioners invoke the supply-side assumptions of fixed technical coefficients, perfectly elastic supplies of inputs and consequently of output. These assumptions may be reasonable in regions where excess capacity and seasonality of labor demand in agriculture allow a highly elastic supply response much of the year. But, where labor supply is tight or short-run capacity constrains, the fixed-price assumption for nontradables becomes more tenuous. For tradables, on the other hand, modelers defend fixed prices on the grounds that they are modeling small, open, regional economies which face a constant, externally determined export price.

The major distinction among fixed-price models lies in their assessment of what constrains tradable output. While the input-output model portrays tradable output as demand-constrained, the semi-input-output and economic base models maintain that tradable output is limited by rigidities on the supply side; that is, by available resources and technology which limit output.

Rather than following the chronological order of their development, from EB to I-O to S-I-O, the following discussion begins with the semi-input-output model. It serves as the reference standard because, within the limitations of fixed-price models, we believe it most closely approximates the reality of third world agriculture. Its supply-constrained view of farm output accords with the frequent technological, land, or water constraints experienced in most developing country agriculture. By comparison, the input-output view that farmers can increase output in unlimited amounts at constant cost strains

credulity in a world where so many go hungry. Consequently, we begin with the most realistic, the semi-input-output model, and then show how the others are variants or special cases of it.

Although the S-I-O and I-O models allow unlimited disaggregation of tradable and nontradable activities, such disaggregation merely adds to the numerical precision of the estimated multipliers rather than to any conceptual understanding.¹ The following description, therefore, proceeds using a two-sector framework, because this greatly simplifies the exposition.

Semi-Input-Output Models (S-I-O)

(i) *Full two-sector S-I-O.* The semi-input-output model developed by Bell and Hazell describes a rural economy in which regional gross output of tradable commodities (T) is fixed by resource or technology constraints at \bar{T} . As in all fixed price models, the S-I-O assumes that nontradable output (N) is perfectly elastic. Its output level is, therefore, determined by local demand. Demand includes household consumption demand for nontradables (H_n), regional producer's intermediate demands for nontradables (D_n), government expenditure on nontradables (G_n), and regional investment demand for nontradables (J_n). Equations (1) through (6) spell out these basic relationships more fully.

- (1) $T = \bar{T}$ tradable output,
- (2) $N = H_n + D_n + \bar{G}_n + \bar{J}_n$ nontradable output,
- (3) $H_n = \gamma_n + \beta_n(Y - L)$ household consumption demand for N ,
- (4) $D_n = a_{nt}\bar{T} + a_{nn}N$ intermediate demand for N ,
- (5) $L = sY$ income leakages,
- (6) $Y = v_nN + v_t\bar{T}$ income definition.

Household consumption expenditure on nontradables increases linearly with income (Y) and depends on the rate of income leakage (L) to savings and taxes, the marginal share of total expenditure directed to nontradables (β_n), and a consumption function intercept (γ_n). Leakages are assumed a constant proportion (s) of total income. Assuming Leontief technology, inter-

¹ The economic base model, on the other hand, does not accommodate disaggregation beyond two sectors.

mediate demand for nontradables remains a constant proportion (a_{ij}) of sectoral gross output, where a_{ij} denotes intermediate deliveries from sector i to sector j . The model further assumes that government and investment expenditures are exogenous at $G_n = \bar{G}_n$ and $J_n = \bar{J}_n$. Finally, the model defines household income (Y) as a constant share of gross output (v_j) in each sector, where $j = T, N$.

Because the model imposes no balance of payments constraint, there is no need to keep track of regional tradable consumption by households or firms. As local consumption varies, exports adjust to equate regional supply and demand, and the model's solution is not affected.

Collecting terms in equations (1) through (6), the model can be written in matrix form, as

$$(7) \quad \begin{bmatrix} 1 - a_{nn} - \beta_n(1-s) \\ -v_n & 1 \end{bmatrix} \begin{bmatrix} N \\ Y \end{bmatrix} = \begin{bmatrix} K_n + a_{nt}\bar{T} \\ v_t\bar{T} \end{bmatrix},$$

where $K_n = \gamma_n + \bar{G}_n + \bar{J}_n$. The region's income depends on \bar{T} and the model parameters, as the solution to (7) indicates:

$$(8) \quad Y = \frac{[(1 - a_{nn})v_t + a_{nt}v_n]\bar{T} + v_nK_n}{1 - a_{nn} - \beta_nv_n(1-s)}.$$

Suppose that investment or technological change in agriculture enables the region to increase its output of tradables. What will be the multiplier impact on the region's income? The answer depends on how the new technology or investment affects the production function for tradables. At a general level, a_{nt} , v_t , and \bar{T} can all be postulated to be functions of θ , a technology shift parameter. In this case, $\partial(\bar{T}v_t)/\partial\theta$ measures the initial change in income in tradables production, and $dY/d\theta$ measures the change in regional total income as a result of the change in technology. The conventional value-added multiplier, which measures the increase in regional value added given a one-unit increase in value added from tradable output, is then obtained from (8), as

$$(9) \quad M = \frac{dY/d\theta}{\partial(\bar{T}v_t)/\partial\theta} = \frac{[(1 - a_{nn})\partial v_t/\partial\theta + v_n\partial a_{nt}/\partial\theta]\bar{T} + [(1 - a_{nn})v_t + a_{nt}v_n]\partial\bar{T}/\partial\theta}{[1 - a_{nn} - \beta_nv_n(1-s)][v_t\partial\bar{T}/\partial\theta + \bar{T}\partial v_t/\partial\theta]}.$$

A much simpler case arises if increases in tradable output do not change the use of inter-

mediate inputs per unit of output. Investments that simply overcome fixed factor constraints could have this effect; examples are land reclamation or additional irrigation in the absence of improved crop varieties. Under these conditions, $\partial v_t/\partial\theta$ and $\partial a_{nt}/\partial\theta$ equal zero, and (9) reduces to

$$(10) \quad M = \frac{1 - a_{nn} + a_{nt}(v_n/v_t)}{1 - a_{nn} - \beta_nv_n(1-s)} = \frac{1}{v_t} \cdot \frac{dY}{d\bar{T}}.$$

To simplify the presentation, we shall consistently adhere to this more restricted type of technological change. This is also consistent with many applied studies that use S-I-O, EB, or I-O models. However, it should be clear that all the models presented can readily be expanded to accommodate more general forms of technological change.

The multiplier captures the combined effects of interindustry linkages and the households' final demand linkages. To separate the two, it is only necessary to solve a variant of the model in which β_n is set equal to zero. Household consumption expenditure then becomes constant and the derived multiplier will be due entirely to production linkages. Modifying (10) in this way and simplifying, the production multiplier (M_p) becomes

$$(11) \quad M_p = \frac{1 - a_{nn} + a_{nt}(v_n/v_t)}{1 - a_{nn}}.$$

(ii) *Hazell's simplified S-I-O.* Hazell has developed a simplified version of the S-I-O model by assuming $a_{nn} = a_{nt} = a_n$ and $v_n = v_t = v$. These assumptions, not unreasonable for many developing country regions, reduce the S-I-O parameter requirements from 6 to 4.² The multiplier in this case becomes

$$(12) \quad M = \frac{1}{1 - a_n - \beta_nv(1-s)},$$

² It reduces the required parameters to three if, as in Hazell's original formulation, the leakage ratio (s) is ignored as well.

while the production linkages alone are

$$(13) \quad M_p = \frac{1}{1 - a_n}.$$

The Economic Base Model (EB)

The economic base (EB) model reduces data requirements still further. It estimates growth multipliers from a single parameter. Because of its extreme simplicity, the EB model has long proven popular, especially among practitioners in urban economics and in regional science (Richardson).

The model refers to a region's tradable sector as its economic base, the output of which, as in the S-I-O model, is constrained at \bar{T} .³ As with all fixed-price models, the economic base (EB) model assumes that nontradable output is perfectly elastic and determined solely by local demand.

The economic base multiplier can be derived in two ways. The first derivation, although less common, highlights the connection between the EB and other fixed-price models. It begins with the income equation in (6) and assumes that nontradable income, $v_n N$, remains a fixed share of total income, Y , that is,

$$(14) \quad v_n N = cY.$$

Total income then becomes

$$Y = v_t \bar{T} + v_n N = v_t \bar{T} + cY = v_t \bar{T} / (1 - c)$$

with a value added multiplier of

$$(15) \quad M = \frac{1}{v_t} \frac{dY}{d\bar{T}} = \frac{1}{1 - c}.$$

When compared to its S-I-O counterparts (10) and (12), it becomes clear that the economic base multiplier is formally identical to the S-I-O. It is a special case of Hazell's S-I-O multiplier arising when

$$(16) \quad c = a_n + \beta_n v (1 - s).$$

Because its single parameter, c , incorporates both

household's final demand and business purchases of intermediates, the economic base model, unlike the S-I-O and I-O models, cannot be used to separate out production from consumption linkages.

A second, more common, derivation of the EB model leads directly to the operational formula preferred by practitioners. From (14) it follows that tradable income $v_t \bar{T}$ is also a fixed share of total income. Its share is $(1 - c)Y$. The ratio (k) of nontradable to tradable income is, therefore,

$$(17) \quad k = v_n N / v_t \bar{T} = c / (1 - c).$$

Thus, the multiplier in (15) is equivalent to

$$(18) \quad M = 1 + k.$$

Practitioners find this version of the multiplier attractive because of its effortless computation once the income data are obtained.

Input-Output Models (I-O)

A closed I-O model with household demand endogenous differs from the S-I-O model in only one respect. It assumes the output of tradables is perfectly elastic and is constrained by export demand alone. Thus, instead of $T = \bar{T}$, the output of tradables is given by

$$(19) \quad T = H_t + D_t + \bar{G}_t + \bar{J}_t + \bar{E}_t,$$

where \bar{E}_t denotes the fixed level of exports. Following an initial spurt in tradable output, the model permits T to increase still further in successive rounds through increased household (H_t) and intermediate demand (D_t), even though \bar{E}_t remains fixed. Because this additional source of growth in regional demand cannot occur in the S-I-O and EB models, the closed I-O model leads inevitably to the largest multipliers.

Defining household (H_t) and intermediate demand (D_t) for tradables analogously to nontradables demand (2) and (4) results in relationships of the following form:

$$(20) \quad H_t = \gamma_t + \beta_t (Y - L)$$

$$(21) \quad D_t = a_n T + a_m N.$$

After combining (19)–(21) and their nontradable counterparts (2)–(4) with the regional income definition (6), the closed I-O model can be written in matrix notation as

³ North, Tiebout, and many others assert that T is constrained by export demand. But this explanation is not consistent with the workings of the model because the model precludes second-round, demand-driven increases in T as a result of an initial increase in its own output. So, in spite of assertions to the contrary, the EB models are formally identical to the S-I-O, and both are founded on the assumption of supply-constrained tradable output.

$$(22) \quad \begin{bmatrix} 1 - a_{nn} & -a_{nt} & -\beta_n(1-s) \\ -a_{tn} & 1 - a_{tt} & -\beta_n(1-s) \\ -v_n & -v_t & 1 \end{bmatrix} \begin{bmatrix} N \\ T \\ Y \end{bmatrix} = \begin{bmatrix} K_n \\ K_t + \bar{E}_t \\ O \end{bmatrix},$$

where $K_n = \gamma_n + \bar{G}_n + \bar{J}_n$ and $K_t = \gamma_t + \bar{G}_t + \bar{J}_t$. Solving this system and calculating the value added multiplier (M) leads to

$$(23) \quad M = \frac{1}{v_t} \frac{dY}{d\bar{E}_t} = [1 - a_{nn} + (v_n/v_t)a_{nt}]/A,$$

where $A = [1 - a_{tt} - \beta_t(1-s)v_t][1 - a_{nn} - \beta_n(1-s)v_n] - [a_{nt} + \beta_n(1-s)v_t][a_{tn} + \beta_t(1-s)v_n]$.

The numerator of (23) is identical to that derived from the S-I-O multiplier (10). The multiplier differs from the S-I-O multiplier only in that the region's demand for tradables now enters the denominator. Moreover, because all the model coefficients are positive and lie between zero and unity, it is clear that A is less than the denominator in (10), and hence the closed I-O multiplier (23) is indeed larger.

By fixing the output of tradables rather than export demand, the S-I-O is equivalent to the closed I-O model when β_t , a_{tn} , and a_{tt} are set equal to zero. The reader can readily verify this by substituting these zero assumptions into (23) to obtain the multiplier in (10).

The classic open I-O multiplier in which household demand is exogenous can now be derived by setting $\beta_t = \beta_n = 0$ in (23). This leads to

$$(24) \quad M_p = \frac{1 - a_{nn} + a_{nt}(v_n/v_t)}{(1 - a_{nn})(1 - a_{tt}) - a_{nt}a_{tn}}.$$

The only difference between (24) and its S-I-O counterpart in (11) lies in the denominator. The two production linkage multipliers are identical when $a_{tt} = a_{tn} = 0$.

Magnitudes

In practice, estimates of agricultural growth multipliers have ranged widely. But, because

practitioners rarely apply different models to the same data set, it is unclear how much of the difference in the multipliers is due to differences in modeling approaches and how much arises because of variations in the underlying economic structure of the regions studied.

To provide consistent comparisons of the different fixed-price models, we apply them all to the same three data sets. Table 1 contains social accounting matrices (SAMs) for three contrasting regions: a low-income, subsistence-oriented agricultural economy (rural Sierra Leone); a middle-income, commercialized agricultural area (the Muda River region in Malaysia); and a high-income, diversified regional economy (Oklahoma state in 1959). In each case, agriculture dominates the regional economy. Yet, input intensity, the level of agroprocessing and trade, and the diversification of nonfarm activity contrast widely among the regions.

The three SAMs provide all but one of the parameters required for the fixed-price models.⁴ The exception, the marginal budget shares (β 's), come from household expenditure studies in Muda (Hazell and Roell) and Sierra Leone (King and Byerlee). The absence of such studies in Oklahoma required educated guesses as to the β 's based on average budget shares available in the SAM. Table 2 displays the fixed-price model parameters obtained for each region.

The empirical results, displayed in table 3, suggest four principal conclusions about fixed-price multipliers.⁵ First, the I-O model produces unrealistically large multipliers. It projects indirect income increments five to ten times as large as its S-I-O counterparts. Consider the difference in Sierra Leone, where the S-I-O model projects \$0.35 indirect income from a \$1 initial injection of tradable income, while the I-O multiplier places the indirect increment ten times as high, at \$3.01. The dramatically different scale arises because the I-O model allows an initial injection of tradable income to stimulate demand not only for nontradables but for more tradables as well. The resulting I-O multipliers

⁴ Because equation (5) assumes that the marginal equals the average leakage ratio, it, too, is available from the SAMs.

⁵ The 1.71 Muda multiplier with 83% consumption linkages reported in table 3 differs from the 1.83 multiplier with 60% consumption linkages reported by Bell, Hazell, and Slade. These differences arise because Bell, Hazell, and Slade (a) simulated the impact of paddy-driven agricultural growth whereas the estimates in table 3 assume an across-the-board increase in the output of all tradables, and (b) adjusted the input-output coefficients for technological change whereas they are held constant in the table 3 estimates.

Table 1. Social Accounting Matrices for Three Regions

	Non-Tradables	Tradables ^a	House-Holds	Capital	Rest of World ^b	Row Sums
Rural Sierra Leone 1974/75						
Nontradables	3.0	2.0	22.8	1.9	.	29.7
Tradables	0.3	1.0	64.0	7.6	27.1	100.0
Households	23.8	95.0	.	.	.	118.8
Capital	.	.	9.5	.	.	9.5
Rest of world ^c	2.6	2.0	22.5	.	.	27.1
	(2.6)	(2.0)	(22.5)	(.)	(.)	(27.1)
Column sums	29.7	100.0	118.8	9.5	27.1	.
Value added	23.8	95.0	.	.	.	118.8
Muda Region, Malaysia 1972, (millions of Malaysian dollars)						
Nontradables	24.1	26.9	205.6	42.6	16.3	315.5
Tradables	20.1	138.2	56.9	.6	180.0	395.8
Households	203.5	169.6	5.9	.	102.1	481.1
Capital	.	.	62.6	.	-2.2	60.4
Rest of world ^c	67.8	61.1	150.1	17.2	78.0	374.2
	(39.8)	(46.1)	(105.1)	(17.2)	(7.1)	(215.3)
Column sums	315.5	395.8	481.1	60.4	374.2	.
Value added	231.4	184.6	.	.	.	416.0
Oklahoma, 1959 (US\$ thou.)						
Nontradables	1,272	909	2,225	144	181	4,731
Tradables	725	1,570	956	195	554	4,000
Households	2,302	1,182	212	165	217	4,078
Capital	156	68	366	.	.	590
Rest of world ^c	276	271	319	86	.	952
	(276)	(271)	(319)	(86)	(.)	(952)
Column sums	4,731	4,000	4,078	590	952	.
Value added	2,458	1,250	.	.	.	3,708

Sources: The Muda SAM was aggregated from table 5.4 in Bell, Hazell, and Slade; nontradables were defined as the aggregate of accounts (19), (21), (22), (24)–(36), (38)–(44); and tradables as the aggregate of sectors (10)–(18), (20), (23), and (37). The Oklahoma SAM was aggregated from table 1 in Little and Doeksen; tradables were defined as the aggregate of the agricultural, agricultural processing, manufacturing, and mining sectors; and nontradables as the aggregate of the government accounts and all remaining productive activities except construction, which was classified as investment and entered in the capital account. The Sierra Leone account was synthesized from a variety of sources, including Matlon et al., Chuta and Liedholm, and King and Byerlee. We do not know the total value of transactions for rural Sierra Leone, hence the SAM is estimated for a gross tradable output of 100.

^a The tradables account does not include direct import of finished goods, which enter the rest-of-the-world account.

^b Includes trade with the nonregional economy, but also institutional accounts (6)–(8) in Muda.

^c Figures in parenthesis are the value of direct imports from outside the region. Any trade and distribution margins are included in the nontradable account.

are excessively large if, as many believe, aggregate agricultural output is supply-constrained in developing countries.

Second, the EB model offers a useful approximation to the S-I-O. Given the ease of computation, it is very attractive, at least as a first approximation to the S-I-O. It produces estimates only slightly worse than Hazell's simplified S-I-O and requires only one parameter instead of four.

Third, comparing linkages across regions, table 3 indicates that the S-I-O multipliers differ substantially, ranging from 1.35 in Sierra Leone to 1.71 in Muda to 2.23 in Oklahoma. Progressively greater input intensity, processing link-

ages, higher income, and consequently increasing consumption diversification into nonfoods account for the substantial spread.

Finally, consumer demand emerges as consistently more important than production linkages, with consumption linkages ranging from 80%–90% of the total multiplier in Muda and Sierra Leone to 60% in Oklahoma. Findings from elsewhere unanimously corroborate the predominance of consumption over production linkages (e.g., Gibb; Bell, Hazell, and Slade; Haggblade and Hazell). Even in low-income regions where consumption linkages are feeble, production linkages tend to be still smaller. This is particularly true in Sub-Saharan Africa, where con-

Table 2. Coefficients for Fixed-Price Models

	Rural Sierra Leone	Muda	Oklahoma
Value-added shares ^a			
v_i	0.95	0.47	0.31
v_n	0.80	0.73	0.52
v	0.92	0.59	0.42
Interindustry coefficients ^a			
a_{ii}	0.01	0.35	0.39
a_{in}	0.01	0.06	0.15
a_{ni}	0.02	0.07	0.23
a_{nn}	0.10	0.08	0.27
a_n	0.04	0.07	0.25
Leakage ratio ^b			
s	0.08	0.13	0.12
Marginal budget shares ^c			
β_i	0.55	0.20	0.25
β_n	0.30	0.50	0.50

^a Derived from the SAMs in table 1.

^b Derived from the SAMs in table 1 by assuming the marginal equals the average leakage ratio.

^c Based on estimates from the following sources: Muda (Hazell and Roell); Sierra Leone (King and Byerlee); Oklahoma projections based on average budget shares. Hazell and Roell report a β_n of 0.4 for Muda, but we increased this to 0.5 to include fresh fruits and vegetables which are produced by "other agriculture," and are classified as a nontradable sector in table 1. Note that $\beta_n + \beta_i$ is less than unity in all cases; the difference is the marginal budget share for direct imports.

sumption linkages are much smaller absolutely than in Asia; however, because production linkages are even lower, the relative share of con-

sumption linkages remains larger (Haggblade and Hazell).

Limitations of the Fixed-Price Models

Fixed-price models are attractive because they are relatively easy to estimate and solve. But the underlying assumption of a perfectly elastic supply of nontradables, common to all fixed-price models, is troublesome. In developing countries, rural nontradables are mostly services or local manufactures produced by small, labor-intensive firms. Given a plentiful supply of labor, a common feature in Asia but also possible in Africa during slack periods in the agricultural calendar, it is reasonable to suppose the short-run supply elasticity of nontradables may be high. Moreover, the low incremental capital-output ratios reported by many researchers for rural nonfarm businesses suggest that even the long-run supply elasticity may be large since firms can expand their capacity with the aid of capital available from retained profits.

But these conditions will not always be met, as the increasing real wages in many successfully growing rural areas attest. In such cases, the supply of nontradables will be upward sloping and part of any agriculture-led increase in demand for nontradables will be dissipated through price inflation. So the multipliers derived from fixed-price models will overestimate the amount of indirect income generated within

Table 3. A Comparison of Fixed Price Income Multipliers in Three Regions

Multiplier (Equation Number)	Rural Sierra Leone	Muda	Oklahoma
Total multiplier			
Semi-input-output (9)	1.35	1.71	2.23
Hazell's simplified S-I-O (12)	1.41	1.49	1.78
Economic base (15)	1.25	2.24	2.98
Input-output (23)	4.01	3.54	6.39
Interindustry multiplier ^a M_p			
Semi-input-output (11)	1.02	1.12	1.53
Hazell's simplified S-I-O (13)	1.04	1.08	1.34
Input-output (24)	1.03	1.73	2.72
Share of multiplier attributable to consumption linkages ^b		(%)	
Semi-input-output	94.3	83.1	56.9
Hazell's simplified S-I-O	90.2	83.7	56.4
Input-output	99.7	71.3	68.1

^a This decomposition is not possible for the economic base model.

^b Calculated as $(M - M_p)/(M - 1) \times 100$.

the regional economy. To correct for this bias requires a price-endogenous model.

A Price-Endogenous (PE) Model

The Model

The proposed model relaxes the two key simplifying assumptions of fixed-price models. First, it allows for increasing-cost nontradable supply by incorporating the two most likely sources of rising nontradable prices. One stems from an upward-sloping labor supply curve facing the region. As the demand for nontradables grows, their increased production increases the regional demand for labor, thereby increasing the nominal wage rate and the average cost of nontradables. The other source of price increase arises from the possibility that capacity constraints and fixed factors, in the short run, lead to diminishing returns in variable inputs and hence to rising average cost as output increases.

Second, to allow less extreme assumptions about input substitution, the proposed price-endogenous model imposes no specific functional form on the production technology. The ensuing empirical applications, however, include a Leontief specification as a special case in order to permit comparison with the fixed-price models.

Following, where possible, the notational conventions used above, the price endogenous model can be written as follows:

for labor in sector j ; π_j , the profit function for sector j ; and θ , a technology shift parameter for production of tradables.

Equations (25) and (26) define market clearing in tradables and nontradables, respectively. Unlike their fixed-price counterparts, equations (2) and (19), market supply now responds to output prices and the nominal wage. Investment and government expenditure remain exogenously determined and, as in the S-I-O model, net exports are a residual. Because there is no balance-of-payments constraint, equation (25) is not part of the simultaneous portion of the solution; it drops out as it did in the S-I-O model.

The labor market representation (27) has no explicit counterpart in the fixed-price models. It is required here to determine the endogenous wage rate. Note that households' willingness to supply labor depends on the real rather than the nominal wage rate.

Regional income (30) now has two components. The first includes profits, that is, returns to fixed factors in each sector. The second includes total wage income, wL_s . In contrast to the fixed-price models where value added (profits plus wage payments) remains a fixed share of gross output in both sectors, the price-endogenous model makes no such presumption.⁶

As with the fixed-price models, we are primarily interested in deriving the income multiplier resulting from a technological shift in the supply of tradables. Note that θ , the technology shift parameter, affects the tradable sector's de-

$$\begin{aligned}
 (25) \quad T(P_t, P_n, w, \theta) &= H_t(P_t, P_n, Y) + D_m(P_t, P_n, w) + D_n(P_t, P_n, w, \theta) \\
 &\quad + \bar{G}_t + \bar{J}_t + \bar{E}_t && \text{tradables,} \\
 (26) \quad N(P_t, P_n, w) &= H_n(P_t, P_n, Y) + D_n(P_t, P_n, w, \theta) + D_m(P_t, P_n, w) \\
 &\quad + \bar{G}_n + \bar{J}_n && \text{nontradables,} \\
 (27) \quad L_s(\bar{w}) &= L_{dt}(P_t, P_n, w, \theta) + L_{dn}(P_t, P_n, w) && \text{labor market,} \\
 (28) \quad \bar{w} &= w/l && \text{real wage,} \\
 (29) \quad l &= P_n^b P_t^{1-b} && \text{cost of living index,} \\
 (30) \quad Y &= \pi_t(P_t, P_n, w, \theta) + \pi_n(P_t, P_n, w) + wL_s && \text{income.}
 \end{aligned}$$

Notation new to the price-endogenous model includes D_{ij} , the intermediate demand by sector j for the output of sector i ; P_t , the price of tradables, presumed as before to be fixed outside the region; P_n , the price of nontradables; the nominal and real wage rates, w and \bar{w} , respectively; l , a cost of living index; b , the share of nontradables in total household income; L_s , the regional labor supply function; L_{dj} , the demand

mand for labor and intermediates as well as its supply function. This formulation provides the flexibility needed to experiment with different types of technological change, although, as with

⁶ In the specific experiments that follow, using Cobb-Douglas technology or Leontief technology with fixed prices, the constancy of value-added shares does remain.

the fixed-price models, we limit ourselves here to forms of investment or technological change that do not affect the use of intermediate goods per unit of tradable output.

Substituting the real wage rate (28), cost-of-living index (29), and income definition (30) into the nontradable (26) and labor market equations (27), totally differentiating with respect to the price of nontradables (P_n), the wage rate (w), the exogenous change in technology (θ), and expressing all variables in percentage change terms (\cdot) leads to⁷

$$(31) \quad \begin{bmatrix} \xi_{P_n}^N - \xi_{P_n}^{H_n} \alpha_n - \xi_{P_n}^{D_{nt}} \alpha_{nt} - \xi_{P_n}^{D_{nn}} \alpha_{nn} & \xi_w^N - \xi_w^{D_{nt}} \alpha_{nt} - \xi_w^{D_{nn}} \alpha_{nn} & 0 & -\xi_Y^{H_n} \alpha_n \\ \xi_{P_n}^{L_{dt}} (1 - \lambda_n) + \xi_{P_n}^{L_{dn}} \lambda_n & \xi_w^{L_{dt}} (1 - \lambda_n) + \xi_w^{L_{dn}} \lambda_n - \xi_w^{L_s} & \xi_{\tilde{w}}^{L_s} & 0 \\ -b & 0 & 1 & 0 \\ -\mu & -d\xi_w^{L_s} & d\xi_w^{L_s} & 1 \end{bmatrix} \begin{bmatrix} \hat{P}_n \\ \hat{w} \\ \hat{f} \\ \hat{Y} \end{bmatrix} = \begin{bmatrix} \frac{\partial D_{nt}}{\partial \theta} \cdot \frac{1}{N} \\ -\frac{\partial L_{dt}}{\partial \theta} \cdot \frac{(1 - \lambda_n)}{L_{dt}} \\ 0 \\ \frac{\partial \pi_t}{\partial \theta} \cdot \frac{1}{Y} \end{bmatrix} d\theta,$$

where new notation includes ξ_j^i , the elasticity of i with respect to j ; shares of gross nontradable output consumed by households ($\alpha_n = H_n/N$) or used as intermediates in nontradable ($\alpha_{nn} = D_{nn}/N$) and tradable production ($\alpha_{nt} = D_{nt}/N$); the proportion of regional labor force employed in nontradables ($\lambda_n = L_{dn}/L_s$); share of regional income spent by households on nontradables ($b = H_n P_n/Y$); wage share of regional income ($d = w L_s/Y$); and the proportion of regional income spent on total final demand for nontradables $\mu = (H_n + \bar{G}_n + \bar{J}_n) P_n/Y$.

Solution of (31) tracks comparative static changes in the two endogenous variables, P_n and w , resulting from the technology shift parameters on the right-hand side. In turn, the changes in labor and nontradable market equilibria generate changes in inflation and regional income according to the bottom two equations in (31). Stability of the system requires that the determinant of the 4×4 coefficient matrix be negative.

Defining the Multiplier

The price-endogenous multiplier is defined as the total increase in real regional income compared to the initial technologically induced increase in tradable income.

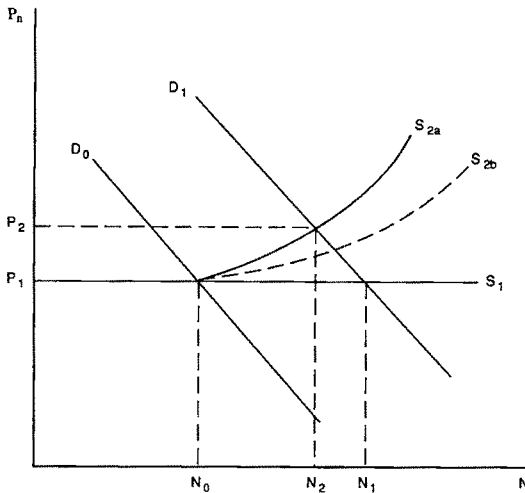
$$(32) \quad M = \frac{\frac{dY}{d\theta} - (N - D_{nn} - D_{nt}) \frac{dP_n}{d\theta}}{\frac{\partial Y}{\partial \theta}}.$$

Unlike the fixed-price models, where no inflation occurs, the price-endogenous model requires a deduction for inflation. The second term in the numerator of (32) represents the increase in the cost of nontradables consumed within the region.⁸ It must be deducted from the total income increase, since we are interested in the change in real rather than nominal income.

Figure 1 pictures the difference between the fixed-price and price-endogenous multipliers. In a fixed-price world, a given increase in technologically induced tradable income will stimulate additional nontradable income equivalent to $(N_1 - N_0)P_1$ times the value added ratio v_n . But when nontradable supply is not perfectly elastic, the same initial increment in tradable income will generate only $(N_2 - N_0)P_1 v_n$ in nontradable earnings. Thus, the upward-sloping nontradable supply dampens the real income

⁷ Derivation of (31) makes use of the following relationships from Sheperd's lemma: $\partial \pi_t / \partial P_n = -D_{nt}$, $\partial \pi_t / \partial P_n = N - D_{nn}$, $\partial \pi_t / \partial w = -L_{dt}$, and $\partial \pi_t / \partial w = -L_{dn}$.

⁸ In computing the real change in regional welfare, we have opted to deduct the entire increase in the cost of nontradables designated for final demand. In addition to households, government and investors consume some nontradables. Because they may or may not be residents of the region, others may wish to deduct only the increased nontradable cost borne by regional residents.



S_1 assumes perfectly elastic supply of labor and nontradables (S-I-O, PE1);

S_{2a} assumes short-run, fixed inputs in nontradables, Leontief technology (PE2a);

S_{2b} assumes short-run, fixed inputs in nontradables, Cobb-Douglas substitutability among intermediates (PE2b)

Figure 1. Graphical comparison of fixed-price and price-endogenous multipliers

multiplier by $(N_2 - N_1)P_1v_n$. Note that calculating the real income multiplier requires deducting the inflation $(P_2 - P_1)N_2 = (N - D_{nn} - D_{nt})dP_n$ from the total nominal income increase in the price-endogenous case.

Economic Interpretation of the Multiplier

The existence of a multiplier depends on drawing unused or underused resources into more productive economic activity. In figure 1, the increase in nontradable output, to either N_1 or N_2 , requires pulling labor, intermediates and in the long run more fixed inputs into nontradable production. To attract resources from other competing uses, including leisure, requires that they earn higher returns in nontradable production.⁹

⁹ The multiplier is defined in terms of income, not utility. A utility-based measure would yield no increase in welfare as a result of a marginal shift of labor out of leisure and into the labor market. Yet the shift from leisure to labor does increase national income. Moreover, welfare will also increase following a finite, as opposed to marginal, increase in labor supply. This will show up as an increase in producer surplus to suppliers of labor.

Formally, the multiplier defined in (32) can be rewritten as follows:¹⁰

$$(33) \quad M = \frac{\frac{\partial \pi_t}{\partial \theta} + w \frac{\partial L_s}{\partial \bar{w}} \frac{d\bar{w}}{d\theta}}{\frac{\partial \pi_t}{\partial \theta} + w \frac{\partial L_{dt}}{\partial \theta}}.$$

It differs from unity to the extent that

$$\frac{\partial L_s}{\partial \bar{w}} \frac{d\bar{w}}{d\theta} \geq \frac{\partial L_{dt}}{\partial \theta}.$$

Thus, the existence of the multiplier depends on drawing some labor into nontradable production. When this happens, the total increase in labor supply (L_s) exceeds the increased labor demand in tradables alone (L_{dt}). Although this formulation centers around the labor market, it implicitly requires that producers similarly draw additional intermediates, and in the long-run fixed factors, into nontradable production.

This derivation yields a surprisingly simple operational means of estimating price-endogenous regional income multipliers resulting from an agricultural project. In addition to project income and employment, an analyst would only require the change in regional employment in nontradables as well as the initial wage rate. Both employment and wage data are relatively easy to obtain, certainly more easily obtainable than changes in total regional income. This does not imply that the multiplier is composed solely of increased wage earnings in nontradables. Profits and returns to fixed factors increase as well. But, through the substitutions from Shepherd's lemma, these portions of the total change in regional income cancel out, because they represent transfers from consumers to producers.

Parameters for Four Experiments

To estimate a price-endogenous income multiplier for a particular region requires data for the twenty-three model parameters in equation (31); twenty parameters in the large matrix and three technology shift parameters on the right-hand side. Of the twenty-three parameters, five of the shares (α_{nn} , α_{nt} , α_n , b , and μ) can be computed

¹⁰ The numerator in (33) is obtained by totally differentiating (30) w.r.t. θ , substituting the results from Shepherd's lemma given in footnote 7, and subtracting the inflation term $(N - D_{nn} - D_{nt})dP_n/d\theta$. The denominator in (33) is obtained by substituting $L_{dt} + L_{dn}$ for L_t in (30), and then taking the partial derivative: $\partial \pi_t / \partial \theta$.

directly from the SAMs. The remaining eighteen must be obtained from outside sources, although economic theory and features of production technology impose a series of restrictions on them. These restrictions and their derivation for the specific functional forms used in this paper are available on request from the authors. Table 4 summarizes the estimated parameter values that apply to each of the three rural regions under review.¹¹

Four variants of the price-endogenous model (*PE1* – *PE4*) permit contrast with their fixed-price relatives by progressively relaxing each of the fixed-price model assumptions. The first, *PE1*, adopts assumptions approximating those of the S-I-O model. It takes supply of both labor and nontradables to be perfectly elastic and specifies Leontief technology in all inputs. The second, *PE2*, introduces an upward-sloping supply of nontradables while maintaining a highly elastic supply of labor. This situation, the one depicted by the supply curve 2a in figure 1, arises in the short run when nontradable producers operate with plants at full capacity and can only increase output at increasing cost. In contrast, *PE3* presumes that labor supply rather than fixed capacity dampens multipliers by raising cost (P_n) via increasing wages. Equivalent to a situation of excess capacity in nontradables, or a long-run time horizon which allows firm entry in nontradables, *PE3* best describes a rural economy during peak agricultural seasons. *PE4* depicts a doubly constrained rural economy in which both the labor supply and nontradable production cost functions slope upwards.

To relax the restrictive S-I-O assumption of fixed-coefficient Leontief production technology, *PE2* and *PE4* include two variants. The first of these, *PE2a* and *PE4a*, retain Leontief technology. But the second pair, *PE2b* and *PE4b*, introduce input substitutability by specifying Cobb-Douglas production technology in all inputs. Expanding substitution possibilities in this way increases nontradable output supply elasticity. By allowing firms to substitute labor for intermediates as the price of nontradable intermediates (P_n) rises, firms are able to increase output at lower cost than would otherwise be possible. The expanded substitution set corresponds to nontradable supply curve 2b in figure 1. It illustrates graphically how allowing for in-

put substitution offsets some of the dampening projected by the price-endogenous multipliers.

Empirical Results

In comparing the series of price-endogenous multipliers with the fixed-price barometer, the S-I-O model, three results are anticipated. First, *PE1*, because it maintains all the perfect elasticity assumptions of the S-I-O, should yield the highest multiplier and it should be equivalent to the S-I-O.¹² Second, the multiplier should be smallest when both labor supply and plant capacity constrain, that is, in *PE4*. In this setting, both wage and price inflation will compound increases in P_n and therefore dampen the multiplier most. Finally, increasing prospects for input substitution should increase nontradable supply elasticity and thus the output response relative to fixed-coefficients technology. Hence, we expect that the multiplier for *PE2b* will be greater than that in *PE2a* and that *PE4b* will exceed *PE4a*. But the ranking of *PE2* and *PE3* cannot be determined a priori. It will depend on the parameters for the specific region in question.

Table 5 bears out these predictions by comparing the multipliers generated in each of the four price-endogenous model experiments. Using data for the three regional economies, Sierra Leone, Muda River region of Malaysia, and Oklahoma, it presents not only multipliers but also the income, wage, and price increases projected in each scenario.

How overoptimistic are the fixed price models? As expected, the answer depends on what is driving the rising price of nontradables. In the worst-case scenario, *PE4a*, the indirect income increment falls dramatically to 20%–40% of the semi-input-output projection (table 5). But after accommodating prospects for input substitution, in *PE4b*, the indirect income increases to 50%–60% of the S-I-O.

In the more likely intermediate situations, *PE2b* and *PE3*, table 5 suggests that price endogeneity will reduce regional indirect income increments to 75%–90% of the S-I-O level. Not surpris-

¹¹ Note that $\partial D_n / \partial \theta \neq 0$ in table 4. Because the commodity demands are written in indirect form in equations (25) and (26), θ must be used to drive increased intermediate demands for tradables production even when the input-output ratios are assumed to remain constant.

¹² For each region, we have calibrated *PE1* to its S-I-O counterpart by making small adjustments in the wage shares of regional income (d). We began with very fragmentary labor share data, often enterprise budgets for selected activities, and sometimes from similar types of agriculture in other countries. Using these as rough approximations generated *PE1* multipliers within 5%–10% of the S-I-O. Because *PE1* approximates the S-I-O when labor supply and output supply are perfectly elastic, this calibration seemed the most plausible procedure for firming up the labor share data.

Table 4. Parameter Values for Price-Endogenous Model

Rural Sierra Leone, 1974/75						Muda River Region Malaysia, 1972						Oklahoma, 1959						
Parameter ^a	PE1	PE2a	PE2b	PE3	PE4a	PE4b	PE1	PE2a	PE2b	PE3	PE4a	PE4b	PE1	PE2a	PE2b	PE3	PE4a	PE4b
Output demand																		
ξ_{Yn}^{bn}	1.09						1.61						1.37					
ξ_{Pn}^{bn}	-1.5						-2.2						-2.2					
Labor supply																		
ξ_w^{Ln}	100	100	100	1	1	1	100	100	100	100	5	5	100	100	100	2	2	2
Shares																		
α_m	0.10						0.08						0.27					
α_n	0.07						0.09						0.19					
α_n	0.77						0.65						0.47					
b	0.19						0.49						0.60					
d	0.69						0.73						0.66					
λ_n	0.38						0.46						0.69					
μ	0.21						0.64						0.69					
Output supply																		
ξ_{Pn}^{Ln}	100	100	100	3.0	3.0	4.0	100	3.0	4.4	100	3.0	4.4	100	2.0	4.6	100	2.0	4.6
ξ_w^{Ln}	-60	-1.8	-2.4	-60	-1.8	-2.4	-55	-1.6	-2.4	-55	-1.6	-2.4	-35	-7	-1.6	-35	-7	-1.6
Labor demand																		
ξ_{Pn}^{Ln}	100	4.0	5.0	100	4	5	100	4.0	5.4	100	4.0	5.4	100	3.0	5.6	100	3.0	5.6
ξ_w^{Ln}	-60	-4.0	-4.0	-60	-4.0	-4.0	-55	-4.0	-4.0	-55	-4.0	-4.0	-35	-3.0	-3.0	-35	-3.0	-3.0
Intermediates																		
ξ_{Pn}^{Ln}	100	3.0	5.0	100	3.0	5.0	100	3.0	5.4	100	3.0	5.4	100	2.0	5.6	100	2.0	5.6
ξ_w^{Ln}	-60	-1.8	-3.0	-60	-1.8	-3.0	-55	-1.6	-3.0	-55	-1.6	-3.0	-35	-7	-2.0	-35	-7	-2.0
Technology shift																		
$\frac{\partial \pi_i}{\partial \pi_i}$	1																	
$\frac{\partial \theta}{\partial \theta}$	Y	0.1					0.1						0.1					
$\frac{\partial D_{ni}}{\partial D_{ni}}$	1																	
$\frac{\partial \theta}{\partial \theta}$	N	0.008					0.022						0.055					

Note: A blank cell takes the value of the closest nonblank cell to its left. Models PE1 and PE3 assume Leontief technology in all inputs, while models PE2 and PE4 consider varying degrees of input substitutability. The "b" variants (PE2b and PE4b) ensure full input substitution via Cobb-Douglas technology in all inputs. The "a" runs (PE2a and PE4a) simulate the fixed-price model assumptions of Leontief technology in intermediates but introduce Cobb-Douglas substitutability in other inputs. Economic theory and these production function specifications impose a series of restrictions on the table parameters. These restrictions and their derivation are available, on request, from the authors.

^a The following 5 parameters are equal to zero and are not listed in the table: ξ_{Pn}^{Ln} , ξ_{Pn}^{Ln} , ξ_{Pn}^{Ln} , ξ_{Pn}^{Ln} , and $\partial L_{ni}/\partial \theta$

Table 5. Price-Endogenous Income Multipliers for Three Regions

Region	PE1	PE2a	PE2b	PE3	PE4a	PE4b
Sierra Leone						
Multiplier	1.35	1.29	1.32	1.26	1.13	1.17
$(M - 1)/(S-I-O - 1)^a$	100	83	91	74	37	49
\hat{Y}	13.5	13.6	13.8	13.1	12.2	12.5
\hat{P}_n	.2	3.5	3.0	2.6	4.0	3.9
\hat{w}	.1	.7	.6	4.2	2.7	3.2
\hat{f}	0	.7	.6	.5	.8	.8
Muda						
Multiplier	1.71	1.40	1.58	1.66	1.30	1.46
$(M - 1)/(S-I-O - 1)^a$	100	56	82	93	42	65
\hat{Y}	17.4	17.8	19.0	17.6	16.8	18.1
\hat{P}_n	.4	6.0	5.0	1.6	6.0	5.4
\hat{w}	.3	3.0	2.6	2.6	3.8	3.9
\hat{f}	.2	3.0	2.5	.8	3.0	2.7
Oklahoma						
Multiplier	2.23	1.46	2.08	2.03	1.22	1.59
$(M - 1)/(S-I-O - 1)^a$	100	37	88	84	18	48
\hat{Y}	22.6	20.7	25.1	22.8	18.0	20.6
\hat{P}_n	.4	8.8	6.3	3.7	8.3	6.8
\hat{w}	.4	5.3	4.0	10.0	6.7	8.5
\hat{f}	.3	5.3	3.8	2.2	5.0	4.1

Note: All units are percentages except the multipliers which are in units of local currency. PE1 is equivalent to the S-I-O model and has Leontief technology and highly elastic supplies of labor and nontradables; PE2a has a highly elastic supply of labor but an upward-sloping supply of nontradables; PE3 has a highly elastic supply of nontradables but an upward-sloping supply of labor; PE4a has upward-sloping supplies of labor and nontradables; and PE2b, PE4b assume Cobb-Douglas technology rather than the Leontief technology of their PE2a and PE4a counterparts. See the text for a more detailed discussion of the assumptions underlying each experiment.

^a Compares the income increment for each multiplier with that from the S-I-O.

ingly, the magnitude of the dampening varies across regions. In the long run, when firm entry allows highly elastic nontradable output responses (PE3), the Muda Region multiplier remains highest, at roughly 90% of the S-I-O level, because of its very elastic labor supply. In contrast, the long-run multipliers in Sierra Leone and Oklahoma fall to 75%–85% of the S-I-O predictions, primarily because of less elastic regional labor supplies.

Conclusions

The existence of agricultural growth multipliers ultimately depends on drawing additional resources into nontradable activity. Hence, the elasticity of nontradable supply is crucial to the size of the multipliers.

Where labor supply is elastic and capacity does not constrain nonfarm activity, fixed-price models are appropriate. Among them, the semi-input-output model offers the best estimate of regional growth linkages. As a rough and ready indicator, the simple one-parameter economic base model offers projections surprisingly close to the semi-input-output model. In contrast, the input-

output model generates unrealistically large multipliers for most rural third world settings.

But price endogeneity may reduce multipliers substantially. Because fixed capacity constraints seem less binding, especially in the long run, labor supply is probably the crucial determinant of the multipliers. If so, our estimates suggest that, as a rough rule of thumb, the Asian multipliers lie in the range of 90% of the levels projected by semi-input-output models, while the African multipliers stand at closer to 75% of the fixed-price estimates.

[Received February 1990; final revision received July 1990.]

References

- Ahmed, C. S., and R. W. Herdt. *A General Equilibrium Analysis of the Effects of Rice Mechanization in the Philippines*. Modelling the Impact of Small Farm Mechanization, Philippine Institute of Development Studies Monograph Series No. 5. Manila, Philippines, 1985.
- Bell, C., P. Hazell, and R. Slade. *Project Evaluation in Regional Perspective: A Study of an Irrigation Project in Northwest Malaysia*. Baltimore MD: Johns Hopkins University Press, 1982.

- Bell, C., and P. Hazell. "Measuring the Indirect Effects of an Agricultural Investment Project on its Surrounding Region." *Amer. J. Agr. Econ.* 62(1980):75-86.
- Chuta, E., and C. Liedholm. *Rural Non-Farm Employment: A Review of the State of the Art*. Dep. Agr. Econ. Rural Develop. Paper No. 4, Michigan State University, 1979.
- Gibb, A. "Agricultural Modernization, Non-Farm Employment and Low Level Urbanization: A Case Study of a Central Luzon Sub-Region." Ph.D. thesis, University of Michigan, 1974.
- Haggblade, S., and P. Hazell. "Agricultural Technology and Farm-NonFarm Growth Linkages." *Agr. Econ.* 3(1989):345-64.
- Haggblade, S., P. Hazell, and J. Brown. "Farm-NonFarm Linkages in Rural SubSaharan Africa." *World Develop.* 17(1989):1173-1201.
- Hazell, P. B. R. "Rural Growth Linkages and Rural Development Strategy." Paper presented at the Fourth European Congress of Agricultural Economics, Kiel, Germany, 3-7 Sep. 1984.
- Hazell, P. B. R., and A. Roell. *Rural Growth Linkages: Household Expenditure Patterns in Malaysia and Nigeria*. Washington DC: International Food Policy Research Institute, Res. Rep. No. 41, 1983.
- Hirschman, A. O. *The Strategy of Economic Development*. New Haven CT: Yale University Press, 1958.
- Hossain, M. *Nature and Impact of the Green Revolution in Bangladesh*. Washington DC: International Food Policy Research Institute Res. Rep. No. 67, 1988.
- Johnston, B. F., and P. Kilby. *Agriculture and Structural Transformation: Economic Strategies in Late Developing Countries*. London: Oxford University Press, 1975.
- King, R. P., and D. Byerlee. *Income Distribution, Consumption Patterns and Consumption Linkages in Rural Sierra Leone*. Dep. Agr. Econ. African Rural Econ. Pap. No. 16, Michigan State University, 1977.
- Krishna, Raj. "Measurement of the Direct and Indirect Employment Effects of Agricultural Growth with Technical Change." *Externalities in the Transformation of Agriculture: Distribution of Benefits and Costs from Development*, ed. Earl O. Heady and Larry R. Whiting. Ames: Iowa State University Press, 1975.
- Little, Charles H., and Gerald A. Doeksen. "Measurement of Leakage by the Use of an Input-Output Model." *Amer. J. Agr. Econ.* 50(1968):921-34.
- Matlon, P., T. Eponou, S. Franzel, D. Byerlee, and D. Baker. *Poor Rural Households, Technical Change and Income Distribution in Developing Countries: Two Case Studies from West Africa*. Dep. Agr. Econ. African Rural Econ. Work. Pap. No. 29, Michigan State University, 1979.
- Mellor, J. W. *The New Economics of Growth: A Strategy for India and the Developing World*. Ithaca NY: Cornell University Press, 1976.
- Mellor, J. W., and B. F. Johnston. "The World Food Equation: Interrelations among Development, Employment, and Food Consumption." *J. Econ. Lit.* 22(1984):524-31.
- Mirakhor, Abbas, and Frank Orazem. "Importance of the Farm Sector to the Economy: A Multiplier Approach." *Amer. J. Agr. Econ.* 50(1968):913-20.
- North, Douglass C. "Location Theory and Regional Economic Growth." *J. Polit. Econ.* 63(1955):243-58.
- Richardson, Harry W. "Input-Output and Economic Base Multipliers: Looking Backward and Looking Forward." *J. Rgnl. Sci.* 25(1985):607-61.
- Tiebout, C. M. "Exports and Regional Growth: A Reply." *J. Polit. Econ.* 64(1956):160-69.

Private Property Rights and Forest Preservation in Karnataka Western Ghats, India

M. G. Bhat and R. G. Huffaker

Areca nut orchard owners have been mulching their orchards with foliage from surrounding government-owned deciduous forests at an average rate that may well deplete foliage within a decade, thereby endangering the long-term viability of the region's most profitable agricultural enterprise. A bioeconomic model of areca nut production is formulated to determine the circumstances under which privatization of forest resources in the hands of orchard owners will stem this rapid depletion. An empirical application suggests that foliage is not currently sufficiently abundant to survive privatization; however, the prospects for preservation can be increased with the imposition of a preprivatization period of regulated use.

Key words: India, optimal extinction, privatization, sustainable agriculture.

This paper considers whether converting hillside forests bordering the areca nut-growing belt of Karnataka Western Ghats in India from government to private ownership by areca nut orchard owners would encourage the sustainable use of forest products in areca nut production and thus stem the rapid depletion occurring under the current use system. The paper represents an application of the recent literature in optimal extinction of renewable resources to study the sustainability of an endemic production process in a part of the world receiving considerable attention by agricultural and resource economists.

We begin the paper by describing the study area and the resource-use problem, then formulate a normative decision model that specifies how a representative orchard owner should harvest foliage from the privately owned section of forest to maximize the present value of areca nut production. We then examine when a sustainable harvest policy is present-value superior to a depletive policy resulting in eventual exhaustion of the foliage stock, under various assumptions about both biological and economic parameters. Finally, by employing economic and environmental data collected from case studies

in the region, we simulate optimizing behavior under private ownership. Assuming that the normative model is valid, the simulations are indicative of the bioeconomic circumstances under which private ownership by areca nut producers is a feasible strategy for preserving forest resources in the study region.

Problem Development

Over the past century, farmers in Karnataka Western Ghats in India have broadened narrow valleys annually by cutting into adjoining hillsides and spreading soil to create terraces. Terraces are privately held and comprised of areca nut orchards (generating the highest economic returns) and crop and paddy lands. Major sources of irrigation water include perennial and seasonal streams flowing through the valleys. The deciduous forests and grasslands covering the hillsides are government-owned and separated into *soppinabetta* (*soppu* = leaf and *betta* = hill) and minor forest.

The *soppinabetta* is essential for the supply of green and dry leaves used to form the farmyard-manure-and-mulch combination that serves to conserve moisture and retard soil erosion in areca nut orchards. The government traditionally has given orchard owners exclusive right in the *soppinabetta* to prune trees for leaves and twigs and to sweep dry leaves from the forest

M. G. Bhat is a graduate research assistant in the Department of Agricultural Economics, University of Tennessee. R. G. Huffaker is an assistant professor, Department of Agricultural Economics, Washington State University, formerly Department of Agricultural Economics, University of Tennessee. Senior authorship is not assigned.

floor. Orchardists also have been allowed to take small timber for fuel and to graze their animals on the grasslands. However, they have not been given the right to chop down large trees or to sell any of the *soppinabetta* products. The government has left orchard owners to divide the *soppinabetta* among themselves by mutual agreement. As a result, each acre of areca nut orchard is supported by about six acres of *soppinabetta*. In contrast, minor forest lands have been reserved for the fuel, leafy matter, and small timber needs of all villagers.

An ecological case study of Karnataka Western Ghats by Gadgil investigated the impact of the utilization of the region's renewable resources (soil, surface and ground water, and plant biomass) on the five major categories of land (crop lands, orchards, *soppinabetta*, minor forest, streams and ponds) for signs of ecological imbalance.¹ Gadgil found that ecological imbalances are related to the pattern of ownership. Government-owned *soppinabetta*, minor forest, and streams and ponds tend to suffer more from the ill effects of imbalanced resource use than privately held orchards and crop land. For example, the success of orchard owners in maintaining soil moisture and inhibiting erosion in areca nut orchards has resulted in an average annual recovery rate of *soppinabetta* foliage biomass (6 tonnes per hectare) that is twice the average annual replenishment rate. Gadgil estimated that this imbalance would deplete the foliage biomass within a decade.

Gadgil briefly discussed two policy options for preserving *soppinabetta* resources. One option is to convert the *soppinabetta* into private ownership by areca nut orchard owners in hopes of achieving the same balanced resource use occurring on privately held lands. The other is to convert the *soppinabetta* into community ownership with community-control systems designed to "ensure serving the objective of preservation of ecological balance with measured benefits to the participants" (p. 386).

A logical first step in deciding which option should be adopted is to determine when private ownership by areca nut producers is a feasible preservationist strategy. Hence, this paper considers the bioeconomic circumstances under which private ownership may lead to ecologically balanced "continuous" use and preserva-

tion of *soppinabetta* resources. A comparison of the relative performance of private ownership versus community ownership and control is beyond the scope of this paper.

The methodology employed depends on the owners' management criterion and the primary productive use they will make of *soppinabetta* resources. Following Clark, the sole owner is assumed "to view the stock as a capital asset; [and thus to manage according to] the standard cost-benefit criterion of maximizing present values of net economic revenues" (pp. 3, 4). Present-value maximization is also a relevant criterion for a community control system; however, the specification of costs and benefits must be expanded to include those external to the sole owner's decision-making environment. This paper's attention is limited to sole-owner behavior because available survey data provide some information on the private net benefits of areca nut production but none on the type or extent of external net benefits (Gadgil; Gadgil, Hedge, and Bhoja Shetty; Hulgol Service Co-operative Society).

The primary productive use areca nut orchard owners will make of *soppinabetta* resources depends on the bundle of property rights that they inherit and on the relative value of alternative outputs. For example, if the bundle includes the right to chop down large trees to sell as lumber, and this is the most lucrative use, then utilization must be studied primarily as a forest rotation problem. Alternatively, if chopping down large trees remains illegal, or areca nut production remains the most valuable use regardless of the legality of harvesting large trees, then utilization must be studied primarily as a conversion of foliage biomass to orchard production. This paper chooses to study the potential for foliage preservation under orchard production because it is the current use and information is available from case studies on the production process and on input/output prices.

A Bioeconomic Model of Areca Nut Production

Consider a representative areca nut producer who mulches one-sixth of an acre of orchard with the foliage biomass generated by an exclusively owned acre of *soppinabetta*. Let $B(t)$ (kilograms per acre, kg/ac) denote foliage density and $R(t)$ (kg/ac/t) the recovery (harvest) rate in year t , where $B, R \geq 0$. Then, following the general

¹ An ecological imbalance occurs when a resource stock decreases over time because it is depleted faster than it can be naturally replenished or when a resource stock accumulates over time because its input rate exceeds its removal rate.

framework of Smith, the dynamics of foliage biomass are governed by

$$(1) \quad \dot{B}(t) = G(B) - R,$$

where $G(B)$ (kg/ac/t) represents the average annual growth (production) of foliage biomass by the forest and is characterized by density dependence and a "minimum viable" biomass stock. Density-dependent growth implies that marginal foliage production tends to increase at low biomass densities but decrease at high densities because of heavier specific competition for vital resources, especially sunlight. The minimum viable biomass stock implies that a forest defoliated on average beyond this level will eventually die out, even if harvesting ceases beforehand. Biomass levels below the minimum viable stock require the forest to refoliate in the same growing season. This, in turn, causes it to begin the next season with low carbohydrate reserves, fewer and less productive leaves, and increased susceptibility to attacks by pathogenic organisms (Wargo, Kramer and Kozlowski).

Density dependence is imposed by requiring $G(B)$ to be strictly concave for $B \geq 0$. The minimum viable stock, \underline{B} , is created by requiring $G(B)$ to have a pair of biomass-axis intercepts, $G(\underline{B}) = G(\bar{B}) = 0$, where $0 < \underline{B} < \bar{B}$. Given an initial biomass level $B_0 > \underline{B}$, the stock tends toward \bar{B} (the carrying capacity) in the absence of recovery. The following generalized logistic equation meets these requirements

$$(2) \quad G(B) = -a_0 + a_1 B - bB^2.$$

A problem with (2) is that it allows a biomass level of zero to decline since $G(0) < 0$. However, this problem can be overcome, while preserving the strict concavity and smoothness of $G(B)$, by imposing the constraint $B \geq 0$ explicitly in optimization (Lewis and Schmalensee, footnote 2). The biomass-axis intercepts are given by

$$(3a) \quad \underline{B} = (1/2b) [a_1 - \sqrt{a_1^2 - 4a_0b}]$$

$$(3b) \quad \bar{B} = (1/2b) [a_1 + \sqrt{a_1^2 - 4a_0b}].$$

Maximum sustained biomass yield occurs at $B_{msy} = a/2b$.

Let $A(R)$ [kg/(1/6)ac/t] represent annual areca nut production on one-sixth of an acre of orchard land. $A(R)$ is assumed to vary solely with recovered foliage biomass; all other inputs are applied in fixed proportions. Positive but diminishing marginal returns to the application of

biomass are assumed; hence $A_R(R) > 0$, and $A_{RR}(R) < 0$, where the subscripts denote differentiation. $A(R)$ has a horizontal asymptote at the maximum periodic production level \bar{A} . The Michaelis equation meets these conditions and is convenient for analytical use:

$$(4) \quad A(R) = \bar{A}(1 - e^{-\alpha R})$$

where α is a positive parameter and q is the fraction of recovery made up solely of leaves (i.e., excluding twigs).

Let p (rupees per kilogram, Rs/kg) denote the constant unit selling price of areca nuts and $C(B)$ (Rs/kg) the cost per unit of biomass recovery. An inverse relationship is assumed between unit recovery costs and biomass to reflect the additional labor required to prune limbs of greater height as foliage biomass decreases, i.e., $C_B(B) < 0$. Moreover, a finite vertical intercept, c_0 , is assumed so that the possibility of extinction as a present-value-maximizing policy is not ruled out a priori. Finally, unit costs are assumed to asymptotically approach zero as biomass increases. An exponential cost function satisfies these characteristics

$$(5) \quad C(B) = c_0 \exp(-c_1 B),$$

where c_1 is a positive parameter.

Periodic profits (Rs/ac/t) generated by employing *soppinabetta* resources in areca nut production are

$$(6) \quad \pi(t) = pA(R) - C(B)R - FC,$$

where FC represents the unavoidable costs of fixed factors associated with areca nut production (mainly pesticide).² The present value of the stream of profits over the interval $0 \leq t \leq T$ is

$$(7) \quad PV = \int_0^T \pi(t)e^{-rt} dt,$$

where r is the real annual discount rate.

The producer's problem is to select a function $R(t)$ and a scalar T that maximizes (7) subject to (1), the nonnegativity of R and B , and a given initial biomass level, $B(0) = B_0$. Following Takayama (theorem 8.C.1), the lagrangian function for this problem is written as

$$(8) \quad L[B, R, \gamma_1, \gamma_2] \\ = H[B, R] + \gamma_1 R + \gamma_2 [G(B) - R],$$

² The present value function does not include the producer's savings generated by using recovered twigs as fuel. Such savings are minor in comparison to profits from areca nut production, and their addition significantly obscures the analytical results.

where $H[B, R]$ is the current-valued hamiltonian measuring the flow of annual net benefits

$$(9) \quad H[B, R] = pA(R) - C(B)R + \lambda[G(B) - R],$$

$\lambda(t)$ is the costate variable measuring the marginal present value of biomass in areca nut production in each year, $\gamma_1(t)$ and $\gamma_2(t)$ are functions associated with the nonnegativity constraints on R and B , and $[G(B) - R]$ is the effective nonnegativity restriction on B .

Solution functions R, λ, B must solve the system of Pontryagin necessary conditions:

$$(10a) \quad L'(R) = pA'(R) - C(B) - \lambda + \gamma_1 - \gamma_2 = 0$$

$$(10b) \quad \dot{\lambda}(t) - r\lambda = -L'(B) = -[-C'(B)R + \lambda G'(B) + \gamma_2 G'(B)]$$

$$(10c) \quad L'(\lambda) = \dot{B} = G(B) - R,$$

the nonnegativity constraint qualifications on R and B (Takayama, theorem 8.C.1; Arrow and Kurtz, chap. 2, proposition 5)

$$(10d) \quad \gamma_1 \geq 0; \quad \gamma_1 R = 0$$

$$(10e) \quad \gamma_2 \geq 0; \quad \gamma_2 B = \gamma_2 [G(B) - R] = 0,$$

the free terminal-time transversality condition (Lewis and Schmalensee, p. 538)

$$(10f) \quad \lim_{t \rightarrow T} e^{-rt} H[B, R] = 0,$$

and nonnegativity restrictions on the terminal biomass level $B(T)$ if the optimal value of T is finite (Takayama, corollary to theorem 8.C.1; Lewis and Schmalensee, p. 538)

$$(10g) \quad \lim_{t \rightarrow T} e^{-rt} \lambda(t) \geq 0$$

$$(10h) \quad \lim_{t \rightarrow T} e^{-rt} \lambda(t) B(t) = 0.$$

When (10a)–(10c) can be satisfied by non-zero levels of R and B [so that γ_1 and γ_2 equal zero by (10d) and (10e), respectively], routine calculation converts (10a)–(10c) into two first-order differential equations in R and B :

$$(11a) \quad \dot{R} = \frac{C'(B)G(B) - [pA'(R) - C(B)][G'(B) - r]}{pA''(R)}$$

$$(11b) \quad \dot{B} = G(B) - R.$$

Recovery Strategies

Recovery strategies can be categorized according to the length of the underlying time horizon. Lewis and Schmalensee refer to infinite horizon strategies as continuous strategies, and finite horizon strategies as abandonment strategies. Infinite horizon strategies are continuous because they asymptotically approach a steady state of (11). Once reached, steady-state biomass and recovery levels are sustained through time by producing both ecological balance (where biomass is recovered at its rate of natural replenishment) and economic balance (where the producer has no further economic incentive to manipulate the recovery rate over time). Alternatively, finite horizon strategies veer away from steady states of (11), and thus ultimately result in the resource being abandoned after some finite time.

This section proceeds by characterizing: (a) the optimal continuous recovery strategy and its domain of feasible initial biomass levels, B_0 ; (b) the optimal abandonment strategy, its domain of feasibility, and when it results in eventual extinction of the stock; and (c) the superiority of one strategy over the other when their domains of feasibility overlap.

Optimal Continuous Recovery Strategy

A continuous recovery strategy, satisfying necessary conditions (10) and asymptotically approaching a steady state of (11), is optimal among all such policies if it satisfies the concavity conditions set out in Arrow and Kurtz (chap. 2, proposition 10). These conditions require that the Hessian matrix of second partial derivatives of the hamiltonian (9) with respect to B and R is negative definite when evaluated at steady state (B^*, R^*). Thus, the first principal minor must be negative:

$$(12) \quad |H_{RR}(R^*)| = pA_{RR}(R^*) < 0,$$

and the second principal minor must be positive:

$$(13) \quad \begin{vmatrix} H_{RR} & H_{RB} \\ H_{BR} & H_{BB} \end{vmatrix} = -pA_{RR}(R^*)[C_{BB}(B^*)R^* - \lambda(B^*, R^*)G_{BB}(B^*)] - C_B(B^*)^2 > 0,$$

where $\lambda(B^*, R^*)$ can be solved for using (10a).

Clearly, (12) holds given the assumed diminishing marginal productivity of foliage biomass in areca nut production. However, (13) is indeterminate, and thus will be numerically evaluated at the steady-state of interest in a later section of this paper.

A phase portrait in (R, B) space is useful for determining the configuration and stability of steady-state solutions of (11).

Foliage biomass zero-change isocline. The foliage biomass zero-change isocline (the locus of B, R combinations that balance growth with depletion and thus hold the biomass stock constant over time) is derived by setting \dot{B} in (11b) equal to zero and solving for R in terms of B ,

$$(14) \quad R(B)|_{\dot{B}=0} = G(B) = -a_0 + a_1 B - bB^2.$$

The biomass zero-change isocline (BI) is the annual growth function (2) and hence shares its properties.

Recovery zero-change isocline. The recovery zero-change isocline (the locus of B, R combinations holding recovery constant over time) is derived by setting $\dot{R} = 0$ in (11a) and solving for R in terms of B ,

$$(15) \quad R(B)|_{\dot{R}=0} = -(1/\alpha q) \ln \left\{ \frac{C_B(B)G(B) + C(B)[G_B(B) - r]}{\alpha q p \bar{A}[G_B(B) - r]} \right\}, \quad \text{or} \\ = -(1/\alpha q) \ln \left\{ \frac{c_0 \exp(c_1 B)[bc_1 B^2 - (a_1 c_1 + 2b)B + [a_0 c_1 + (a_1 - r)]]}{\alpha q p \bar{A}(a_1 - 2bB - r)} \right\},$$

in terms of (2), (4), and (5). The recovery zero-change isocline (RI) is considerably more complicated to characterize than the biomass isocline. This is the result of the restricted domain of the logarithmic function and the discontinuities that result when the numerator and denominator in its argument (bracketed term) are identically zero.

In characterizing RI , denote the numerator and denominator in the logarithmic term of (15) as $\text{Num}(B)$ and $\text{Den}(B)$. Let B_+^n and B_-^n ($B_+^n > B_-^n$) represent the roots of the quadratic equation in $\text{Num}(B)$ derived from adding and subtracting the radical in the quadratic formula, respectively. Let $B^r = (a_1 - r)/2b$ denote the root of $\text{Den}(B)$. B^r represents the "golden-rule" stock level prevalent in capital-based models that equates the marginal productivity of the biomass

stock (the biological rate of interest) with the financial rate of interest, i.e., $G'(B^r) = r$.

Assume that the intrinsic growth rate of biomass is greater than the discount rate, i.e., $a_1 - r > 0$. This appears reasonable for the study area over a wide range of circumstances (see table 1) and decreases the enumeration of possible cases. Given this assumption, application of Descartes' Rule of Signs (Albert, p. 171) to $\text{Num}(B)$ implies that B_+^n and B_-^n are either both positive real roots or a complex conjugate pair; B^r is clearly positive. When B_+^n and B_-^n are real, RI can be shown to have vertical asymptotes approaching positive infinity at B_+^n and B_-^n , and a vertical asymptote approaching negative infinity at B^r . Figure 1 shows how the configuration of RI depends on the magnitude of B^r relative to B_+^n and B_-^n .³

The position of RI in phase space depends critically on the magnitudes of the limiting marginal profitability of recovering the final unit of biomass (referred to hereafter as limiting marginal profitability)

$$(16) \quad \lim_{(R,B) \rightarrow (0,0)} \partial \pi(t) / \partial R = \alpha q p \bar{A} - c_0,$$

and on a_0 , which is directly related to the minimum viable biomass stock, \underline{B} by (3a). In figure 1a-c, a *ceteris paribus* increase in limiting mar-

ginal profitability drives both the leftward and rightward portions of RI upward. A *ceteris paribus* increase in \underline{B} also drives the rightward portion upward, but drives the leftward portion downward.

Steady-state solutions. Steady-state solutions of (11) occur at the intersections of the biomass and recovery zero-change isoclines. Figure 2 shows a phase diagram, associated with the RI depicted in figure 1b, depicting a three-steady-state configuration. Clearly, other three-steady-state configurations are possible if the RI 's in

³ When B_+^n and B_-^n are a complex conjugate pair, the rightward portion of RI is no longer defined, and the leftward portion is similar to that in figure 1a.

Table 1. Baseline Parameter Values

Symbol	Meaning	Units	Value
a_1	intrinsic annual growth rate of biomass	yr^{-1}	$(a_0 = 500)$ 0.158 $(a_0 = 5,000)$ 0.583
b	density dependence of biomass growth	$(\text{kg}/\text{ac})^{-1}\text{yr}^{-1}$	$(a_0 = 500)$ $6.5E-7$ $(a_0 = 5,000)$ $3.1E-6$
\bar{A}	max. annual areca nut production level	$\text{kg}/(1/6)\text{ac}/\text{yr}$	300
α	production parameter		0.00039
q	fraction of recovered biomass in greenleaf		0.9
p	unit return of areca nuts	Rs/kg	14.81
c_1	unit cost parameter		$(c_0 = 0.3)$ $1.3E-5$ $(c_0 = 1.7)$ $1.7E-4$
r	real annual discount rate		0.047

figures 1a,c are used. In addition, figure 2 bifurcates to a one-steady-state configuration if, for example, limiting marginal profitability is sufficiently large relative to other parameters to position the leftward portion of RI above BI . However, an analysis based on the three-steady-state configuration in figure 2 suffices to elucidate the general issues.

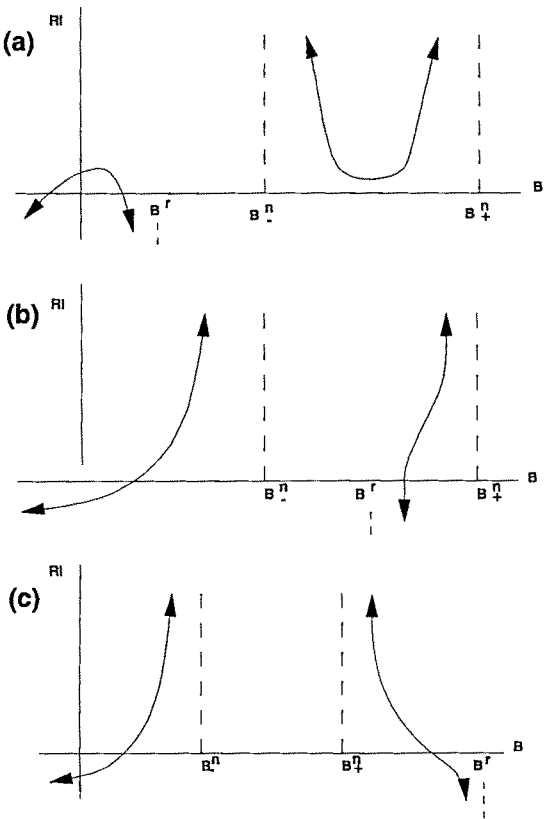


Figure 1. The recovery zero-change isocline, RI : (a) $B^r < B^- < B^+$; (b) $B^- < B^r < B^+$; (c) $B^- < B^+ < B^r$

Direction of motion. The arrows in figure 2 depict the direction of motion away from steady state. Since $\dot{B}'(R) < [0 \text{ by (11b)}]$, biomass densities tend to decrease over time for recovery rates above BI and increase for those below. Recovery rates can be shown [by signing $\dot{R}'(R)$ in the following intervals] to decrease for biomass levels to the right of the rightward portion of the recovery zero-change isocline; increase for biomass levels between the rightward and leftward portions; and decrease for biomass levels to the left of the leftward portion.

Local stability properties. Local stability properties of the steady states in figure 2 can be examined by linearizing system (11) in the neighborhood of each steady state. The Jacobian matrix of this linearized system is J with a characteristic polynomial

(17) $CP(E) = E^2 - Tr(J)E + Det(J) = 0,$

where $Tr(J) = \dot{R}'(R) + \dot{B}'(B)$ is the trace and $Det(J) = \dot{R}'(R)\dot{B}'(B) - \dot{B}'(R)\dot{R}'(B)$ is the determinant of J . $Det(J)$ can be shown to be negative for steady-state biomass stocks greater than the maximum sustained yield level, B_{msy} . This implies that $CP(E)$ has roots (eigenvalues) of opposite sign, which in turn indicates that a steady state occurring to the right of B_{msy} is a saddle-point equilibrium.

Alternatively, $Det(J)$ is indeterminate for steady-state stocks to the left of B_{msy} ; hence, stability properties cannot be guaranteed by this procedure short of numerical computation of the associated eigenvalues. However, stability properties can be deduced from the direction of motion of the system (as depicted by the arrows in fig. 2). Motion around the outer steady states (denoted as SS_L and SS_H) is consistent with a saddle-point equilibrium, whether or not SS_H is

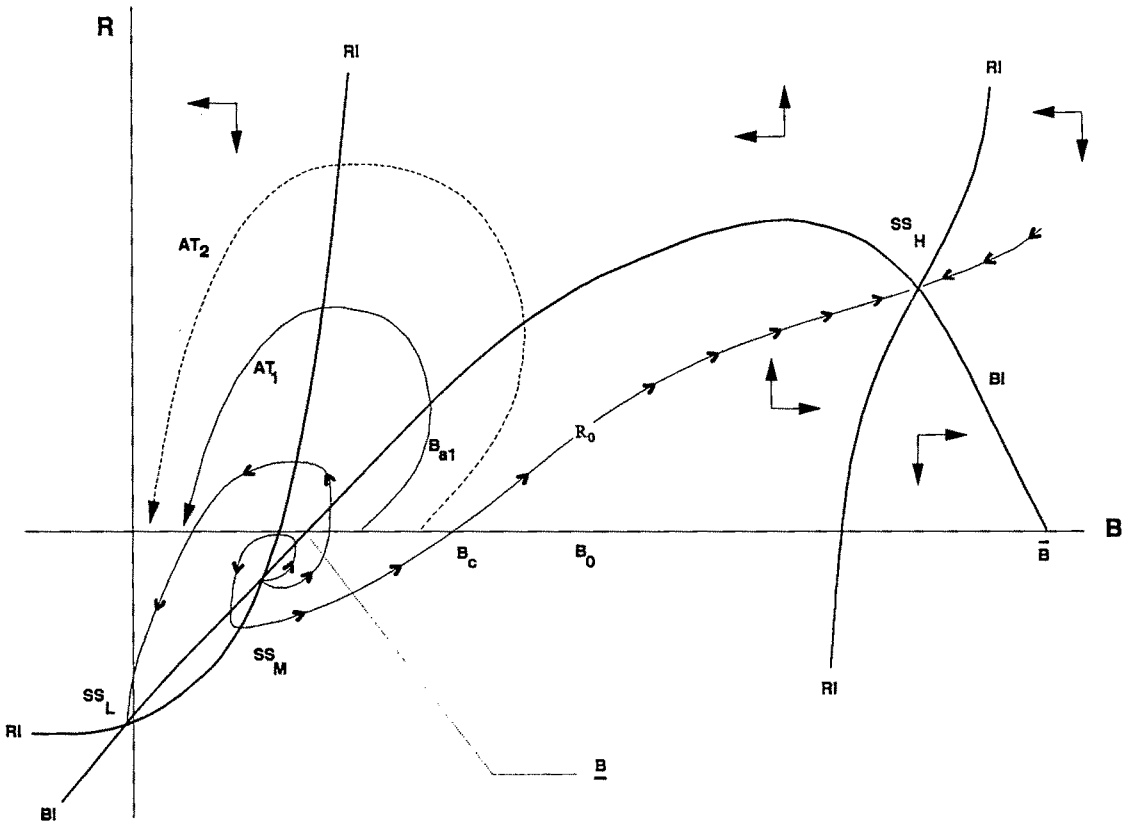


Figure 2. Phase diagram with three-steady-state configuration

to the right of B_{msy} . Motion around the middle-steady state, SS_M , is consistent with either an unstable focus (eigenvalues are a complex conjugate pair with positive real parts) or an unstable node (eigenvalues are positive real numbers).⁴ In either case, two convergent separatrices (arrowed trajectories) emanate from SS_M and asymptotically approach their respective outer steady states.

In figure 2 the optimal continuous recovery strategy [assuming that sufficient condition (13) is met] is to select initially the recovery rate $R(0) = R_0$ lying directly above $B(0) = B_0$ on the separatrix converging upward toward the positive-biomass steady state SS_H .⁵ The upward conver-

gent separatrix approaching SS_H is consistent with the nonnegativity restrictions on B and R for $B_0 \in [B_c, B_H]$, where B_c is the separatrix' largest biomass-axis intercept. The downward convergent separatrix approaching SS_H satisfies nonnegativity for $B_0 \in [B_H, \infty)$. Hence, the optimal continuous strategy is feasible for $B_0 \in [B_c, \infty)$.

Optimal Abandonment Strategy

The optimal abandonment strategy is selected from among paths satisfying (11) but veering away from steady state. These paths imply abandonment within a finite time, T , by crossing either the biomass axis (where there is no further economic incentive to recover biomass) or the recovery axis (where recovery necessarily ceases because the biomass is exhausted). The optimal path among these satisfies the free terminal-time transversality condition (10f) requiring that the Hamiltonian (9) must approach zero as t approaches T , and the nonnegativity restric-

⁴ Davidson and Harris (appendix a) rule out the possibility of closed cycles as optimal trajectories in optimal control problems with exponential discounting and one state variable and one costate variable.

⁵ In the case where the lower-biomass saddle point, B_L , is positive and satisfies the sufficient conditions, and both B_L and B_H are approachable from B_0 , the extramarginal analysis set out in Davidson and Harris may be used to select the convergent separatrix representing the optimal continuous policy.

tions on the terminal biomass level $B(T) = B_T$. The nonnegativity restrictions require the costate variable to approach a nonnegative limit (10g), and the product of the costate and state variables to approach zero (10h), as t approaches T . Condition (10h) implies one of the following three cases: (1) B_T is positive and $\lambda(T) = \lambda_T$ is zero; (2) Both B_T and λ_T are zero; or (3) B_T is zero and λ_T is positive.

Case 1. $B_T > 0$, $\lambda_T = 0$. The optimal levels of B_T and R_T must satisfy

$$(18a) \quad \lambda(T) = pA'(R_T) - C(B_T) = 0$$

$$(18b) \quad H(T) = pA(R_T) - C(B_T)R_T = 0,$$

where $\lambda(T)$ is derived from (10a). Condition (18a) requires marginal revenue of recovery, $pA'(R_T)$, to equal marginal cost, $C(B_T)$, at the optimal terminal recovery and biomass levels. Condition (18b) requires average revenue of recovery, $pA(R_T)/R_T$, to equal average cost, $C(B_T)$. Both conditions are met when marginal revenue equals average revenue. Given the concavity of $A(R)$, marginal revenue lies below average revenue for $R > 0$, hence equality occurs only when $R_T = 0$. Solving (18a) for optimal B_T when $R_T = 0$ yields

$$(19) \quad B_T = -(1/c_1) \ln(paq\bar{A}/c_0),$$

which is restricted to be positive in Case 1, implying that $paq\bar{A} < c_0$. The optimal terminal stock condition (19) requires that B_T increase (thereby promoting conservation of the stock) as the limiting marginal profitability of recovering the last unit of biomass (16) becomes increasingly negative.

In Case 1, the producer abandons the *soppinabetta* before foliage is completely exhausted because recovery is no longer economic. Abandonment results in eventual extinction if the stock is driven to a terminal level to the left of the minimum viable level (3a), i.e., $B_T < \underline{B}$. Setting $B_T < \underline{B}$ in (19) implies eventual extinction when limiting marginal recovery costs, c_0 , are less than a critical level that increases with the minimum viable stock,

$$(20) \quad c_0 < paq\bar{A} \exp(c_1 \underline{B}).$$

Alternatively, Cases 2 and 3 above both stipulate a priori that the optimal abandonment policy will result in extinction because B_T is set to zero. The difference between the two cases is in the optimal terminal recovery rate, R_T .

Case 2. $B_T = \lambda_T = 0$. Conditions (18a) and (18b) (evaluated at $B_T = 0$) continue to be relevant and require that marginal revenue equals average revenue at R_T ; hence, R_T continues to be identically zero. This is economically reasonable because specifying that $B_T = \lambda_T = 0$ imposes the restriction that the limiting marginal profitability of recovering the last unit equals zero [by (19)]. Hence, the terminal recovery rate is zero because the producer would have no economic incentive to harvest another unit even if it could be produced.

Case 3. $B_T = 0$, $\lambda_T > 0$. The optimal level of R_T must now satisfy

$$(21a) \quad \lambda(T) = pA'(R_T) - C(B_T = 0) > 0$$

$$(21b) \quad H(T) = pA(R_T) - C(B_T = 0)R_T + \lambda_T[G(B_T = 0) - R_T] = 0.$$

Substituting in (2), (4), (5) for $G(B)$, $A(R)$, $C(B)$, respectively, and λ_T from (10a), into (21b), yields after some manipulation,

$$(21c) \quad H(T) = p\bar{A} + c_0a_0 - p\bar{A} \exp(\alpha q R_T) [1 + \alpha q(a_0 + R_T)] = 0.$$

Condition (21c) is convex in R , and under the specifications of Case 3, has two roots of opposite sign. The positive root is the optimal terminal recovery level R_T . In contrast to Case 2, the optimal terminal recovery rate is positive because recovery remains profitable as biomass reaches extinction [$\lambda_T > 0$ implies that the limiting marginal profitability of recovering the last unit is positive].

Figure 2 shows the unique abandonment trajectory, AT_1 , associated with the optimal terminal pair of recovery rate and biomass stock in Case 1. Let the largest biomass level lying on AT_1 be denoted as B_{a1} . Then, the optimal abandonment strategy is feasible for initial biomass levels in the interval $B_0 \in [B_T, B_{a1}]$. The producer engages this strategy by initially selecting the unique recovery rate, R_0 , lying directly above the initial biomass level, B_0 , on AT_1 .

Optimal Recovery Strategies

Consider the problem of whether to select an optimal continuous strategy or an optimal abandonment strategy over the entire domain of initial biomass levels. Assume first that the feasible domains for each strategy do not overlap.

This occurs with optimal abandonment trajectory AT_1 in figure 2, which is drawn so that AT_1 does not lie anywhere above the upward convergent separatrix for $B > B_c$. Each strategy is the overall optimal strategy within its feasible domain, and the superiority of one over the other is not an issue.

Neither strategy is feasible for initial biomass levels in the intervals $B_0 \in (0, B_T)$ and $B_0 \in (B_{a1}, B_c)$. The only feasible choice from such initial stocks is among suboptimal abandonment trajectories that eventually intersect either the recovery axis (given relatively "large" initial recovery rates) or the biomass axis to the left of \underline{B} (given relatively "small" initial recovery rates) and thus result in eventual extinction. The choice among these trajectories must be made by computing the associated present values.

In summary, when the feasible domains of the optimal strategies do not overlap, the producer will select a continuous strategy leading to ecological balanced use if $B_0 \geq B_c$, or some type of abandonment strategy resulting in eventual extinction if $B_0 < B_c$.

Assume, alternatively, that the feasible domains of the optimal abandonment and continuous strategies overlap, thereby raising the issue of which strategy is superior in the overlapping interval. The optimal abandonment trajectory is now AT_2 (dashed path in fig. 2), which lies above the upward convergent separatrix for a small interval of $B_0 > B_c$. Cropper (proposition 1) shows in a generalized resource extraction model that, when $0 \leq B_T \leq \underline{B}$ (as for AT_2), a $B_0 > \underline{B}$ exists such that the optimal abandonment strategy is superior to the optimal continuous strategy and extinction is therefore optimal. Denote this critical initial stock as B_{cr} . Then, for $B_0 \leq B_{cr}$ an abandonment strategy is overall optimal, and for $B_0 > B_{cr}$ the optimal continuous strategy is overall optimal. The economic rationale is that growth is sufficiently slow for $B_0 < B_{cr}$ that it does not pay to build the stock up toward SS_H at the expense of foregoing the net benefits of driving the stock to extinction.

Simulated Areca Nut/*Soppinabetta* Dynamics

This section uses available case study information to generate numerical simulations of areca nut/*soppinabetta* dynamics in (11). The goal is to identify circumstances under which (a) a representative present-value-maximizing producer in the study area favors abandonment and even-

tual extinction over a continuous recovery strategy; and, hence, (b) granting ownership of the *soppinabetta* to areca nut producers is not a feasible strategy for preserving the resource.

The numerical simulations rely mainly on parameter values calculated from information on economic and ecological conditions provided in case studies of the region found in Gadgil, Gadgil, Hedge, and Bhoja Shetty; and Hulgol Service Cooperative Society. Table 1 summarizes the baseline parameter values derived below.

Areca nut yield function parameters. Gadgil reported an average annual yield, $A(R)$, of 2.56 tonnes per hectare [173.085 kg/(1/6)ac] with the application of an average annual recovery, R , of 6 tonnes per hectare (2,434.02 kg/ac). Our own consultations with area producers resulted in the following values: $\bar{A} = 300$ kg/(1/6)ac/yr, and $q = 0.9$ (the fraction of recovery in foliage). Inserting these values into areca nut yield function (4) gives

$$(22) \quad 173.085 = 300(1 - e^{-\alpha(.9)(2434.02)}),$$

which can be solved for $\alpha = 0.00039$.

Areca nut unit return and discount rate. Gadgil reported that the average annual areca nut yield of 2.56 tonnes per hectare [173.085/(1/6)ac] generated an annual income of approximately rupees 38,000 per hectare [Rs 2,564.1/(1/6)ac]. Hence, the unit return is calculated as $p = 2,564.1/173.085 = \text{Rs } 14.81/\text{kg}$. The assumed nominal discount rate, $r_{\text{nominal}} = 0.1$, is the annual fixed deposit rate offered by commercial banks (Dena Bank, interest rate schedule, 1987). The real rate, $r = 0.047$, is calculated by netting out the annual percentage inflation rate (5.3%) over the period 1986–87 (Quarterly Economic Report of the Indian Institute of Public Opinion).

Plant growth parameters. Gadgil reported an average standing biomass of 28 tonnes per hectare (11,336 kg/ac) in his study area compared to 400 tonnes per hectare (161,944 kg/ac) in a good moist deciduous forest. The annual productivity of the above two types of forest were 3 tonnes per hectare (1,214.58 kg/ac) and 20 tonnes per hectare (8,097.2 kg/ac), respectively. These values are inserted into $G(B)$ in (2) to form the following two simultaneous equations:

$$(23a) \quad 1,214.58 = -a_0 + a_1(11,336.1) \\ - b(11,336.1^2)$$

$$(23b) \quad 8,097.20 = -a_0 + a_1(161,944) \\ - b(161,944^2).$$

Applying Cramer's Rule to solve this system for a_1 and b in terms of a_0 yields

$$(24a) \quad a_1 = \frac{a_0(11,336.1^2 - 161,944^2) + (8097.2)(11,336.1^2) - (1214.58)(161,944^2)}{(161,944)(11,336.1^2) - (11,336.1)(161,944^2)}$$

$$(24b) \quad b = \frac{a_0(11,336.1 - 161,944) + (8,097.2)(11,336.12) - (1,214.58)(161,944)}{(161,944)(11,336.1^2) - (11,336.1)(161,944^2)},$$

where a_0 is discussed below.

Recovery and cultivation unit costs. A case study by the Hulgol Service Co-operative Society reported that an average of 42 laborer days per acre were employed annually to recover foliage biomass from the *soppinabetta*. This results in a total recovery cost of rupees 630 per acre when multiplied by the average payment per laborer day (Rs 15 per laborer) (obtained in our own consultations with producers). Assuming that this total recovery cost is associated with Gadgil's reported figures on the average annual recovery rate of 6 tonnes per hectare (2,434.02 kg/acre) at the average biomass level of 28 tonnes per hectare (11,336.1 kg/acre), the unit cost of recovery is calculated as rupees 0.26 per kilogram (630/2,434.02). Hence, one point on the unit cost curve is $C(B_1) = 0.26$ and $B_1 = 11,336.1$. Unit cost function (5) can be used to solve for c_1 in terms of c_0 at this point:

$$(25) \quad c_1 = -(1/11,336.1) \ln(0.26/c_0),$$

where c_0 is discussed below.

Parameters a_0 and c_0 . Recall that a_0 determines the minimum viable stock level, \underline{B} , in (3a), c_0 is the limiting marginal cost of harvesting the final unit of foliage biomass in (16), and that both parameters are critical in determining whether the optimal abandonment strategy derived in Case 1 above (the only case not specifying extinction a priori) results in extinction by (20).

Unfortunately, information was not found to calculate these key parameters. The best that can be done currently is to simulate system (11) over

wide intervals for a_0 and c_0 to determine a range of dynamic behavior. Hence, four simulations were run alternating between high and low values for each. Growth parameter a_0 was set so that the minimum viable stock, \underline{B} , alternated between a "low" of approximately 3,000 kilograms per acre and a "high" of 9,000 kilograms per acre. Cost parameter c_0 alternated between a "low" level (0.3 Rs/kg) not satisfying the ex-

inction condition (20) for Case 1 above, and a "high" level (1.70 Rs/kg) satisfying it.

Each simulation recalculated growth parameters a_1 (24a) and b (24b), and cost parameter c_1 (25) so that the growth (2) and unit cost (5) functions continued to run through their respective observed data points, thereby maintaining use of this information (table 1). Increasing a_0 under these conditions squeezes the biomass zero-change isocline inward so that the new curve has a larger maximum sustained yield at a smaller stock level, a larger minimum viable population level, and a smaller carrying capacity (fig. 3a). Increasing c_0 (from c_{01} to c_{02}) under these conditions causes the unit cost curve to pivot around the observed data point with unit costs lying above (below) the old curve for stocks smaller (larger) than the observed point (fig. 3b).

Simulation Results and Discussion

Figures 4a-d show the numerically generated phase diagrams associated with each simulation.⁶ The inset to figure 4c enlarges the portion containing the optimal abandonment trajectory in order to highlight its location relative to the optimal continuous recovery trajectory. Table 2 summarizes steady-state stock and recovery levels along with eigenvalues and sufficient conditions (12) and (13) calculated for each steady state. Because figure 4 and table 2 are fairly comprehensive, the paper forgoes discussion of each simulation separately in favor of a general discussion summarizing the impacts of alternat-

⁶ The phase diagrams were generated by the BDPHS subroutine of Dynamical Software, Dynamical Systems, Tucson AZ.

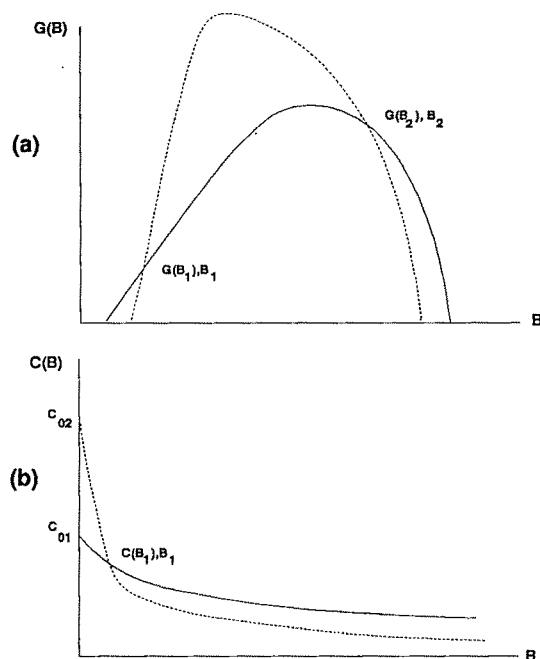


Figure 3. (a) Impact on $G(B)$ of increase in a_0 ; (b) impact on $C(B)$ of increase in c_0 (dashed curves are post-increase curves)

ing between high and low levels for c_0 and B under the previously stated conditions.

Three of the simulations (figs. 4a, b, d) produce the single positive saddle point variation of figure 2 wherein the leftward portion of the recovery isocline (RI) lies above the biomass isocline (BI). Alternatively, the high c_0 –low B simulation (fig. 4c) produces the three-steady-state configuration shown in figure 2. Two steady states are added because high c_0 positions RI further to the right [by decreasing limiting mar-

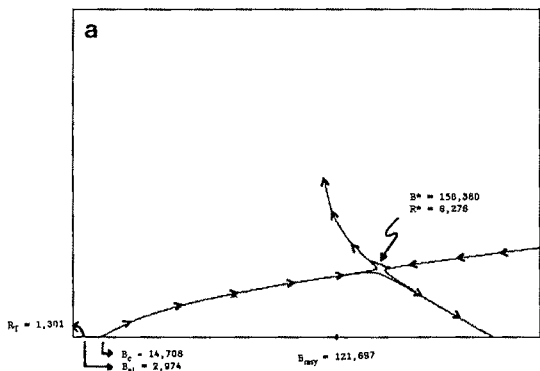


Figure 4a. Phase diagram for low c_0 and low B

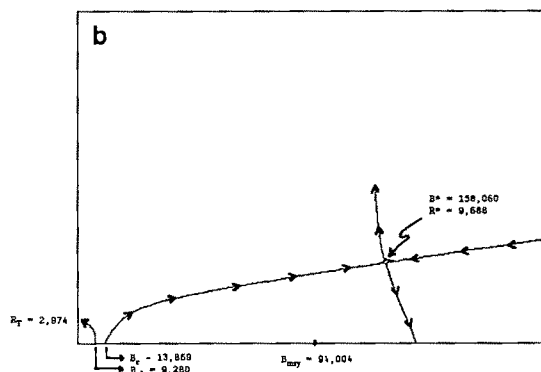


Figure 4b. Phase diagram for low c_0 and high B

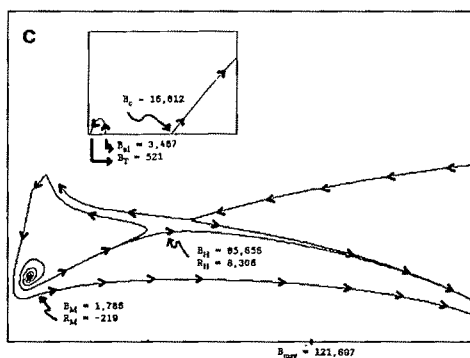


Figure 4c. Phase diagram for high c_0 and low B . The inset enlarges the diagram in the area of the unstable focus B_M, R_M

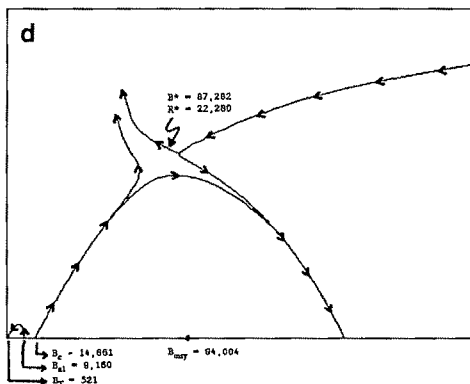


Figure 4d. Phase diagram for high c_0 and high B

ginal profitability (16)] and low B positions BI further to the left. The middle steady state, SS_M , is an unstable focus. The higher-biomass steady state, SS_H , is the only approachable equilibrium meeting the nonnegativity conditions on B and R .

Table 2. Steady States, Stability, and Sufficient Conditions

	Low \underline{B}	High \underline{B}
low c_0	$(B^*, R^*) = (158380, 8276)$ Eigenvalues = $-.083, .130$ First principal minor = $-3.0E-5$ Second principal minor = .43	$(B^*, R^*) = (158060, 9688)$ Eigenvalues = $-.433, .480$ First principal minor = $-1.8E-5$ Second principal minor = .06
high c_0	$(B_H, R_H) = (85656, 8306)$ Eigenvalues = $-.042, .089$ First principal minor = $-3.0E-5$ Second principal minor = .43 $(B_M, R_M) = (1786, -219)$ Eigenvalues = $.024 \pm .30i$ $(B_L, R_L) = (-513945, -254190)$ Eigenvalues = $-.502, .549$ First principal minor = $-3.1E-35$ Second principal minor = $-2.7E80$	$(B^*, R^*) = (87282, 22280)$ Eigenvalues = $-.119, .167$ First principal minor = $-2.2E-7$ Second principal minor = $2.4E-5$

The saddle points in figures 4a, b, d, and SS_H in figure 4c all meet sufficient conditions (12) and (13). Hence, the upward convergent separatrix associated with each is the optimal continuous strategy for $B_0 \in [B_c, B_H]$, and the downward convergent separatrix is the same for $B_0 \in [B_H, \infty)$.⁷ Interestingly, the initial feasible biomass stock on the upward convergent separatrix, B_c , varies little with changes in either c_0 or \underline{B} . The intervals containing B_c ranged from (13868, 13869) for low c_0 /high \underline{B} (fig. 4b) to (16811, 16812) for high c_0 /low \underline{B} (fig. 4c).⁸ Moreover, the feasible domains of the optimal continuous strategy and the optimal abandonment strategy do not overlap in any of the simulations. Hence, each strategy is overall optimal in its feasible domain.

The major impacts of increasing c_0 , while holding \underline{B} constant at either a low or high level, are mostly felt in the equilibrium stock, B^* , and the optimal terminal stock, B_T , and recovery rate, R_T . For low c_0 (figs. 4a, b), B^* is to the right of the maximum sustained yield stock, B_{msy} , (thereby guaranteeing a saddle-point equilibrium), and is a large percentage of the carrying capacity, \bar{B} (67% for low \underline{B} and 88% for high \underline{B}). However, for high c_0 (figs. 4c, d), B^* is to the left of B_{msy} (calculated eigenvalues indicate

that B^* remains a saddle-point equilibrium), and is a much smaller percentage of \bar{B} (36% for low \underline{B} and 49% for high \underline{B}). The reason that B^* moves to the left is that increased c_0 causes the new unit cost curve to lie below the old curve for biomass levels to the right of the observed data point ($B = 11,336.1$). This decreases the recovery cost savings of maintaining the larger equilibrium stocks in figures 4a, b.

Consider now the impact of c_0 on the optimal terminal point (B_T, R_T). For low c_0 (figs. 4a, b), limiting marginal profitability (16) is positive, hence the simulations follow Case 3 above where $B_T = 0$ and $R_T > 0$. For high c_0 (figs. 4c, d), limiting marginal profitability (16) is negative, hence the simulations follow Case 1 above where $B_T > 0$ and $R_T = 0$.

The major impacts of increasing \underline{B} (while holding c_0 constant at either a high or low level) occur in the steady state recovery rate, R^* , and the largest biomass stock on the optimal abandonment trajectory, B_{a1} . When \underline{B} is low (figs. 4a, c), steady-state recovery rates are similar for low c_0 ($R^* = 8,276$ kg/ac/t) and high c_0 ($R^* = 8,306$ kg/ac/t). However, when \underline{B} is high (figs. 4b, d), steady-state recovery rates increase and are markedly different for low c_0 ($R^* = 9,688$ kg/ac/t) and high c_0 ($R^* = 22,280$ kg/ac/t). Steady-state recovery rates increase because the associated steady-state biomass levels are situated to the left of the far-right observed data point ($B_2 = 161,944$ kg/ac) and the biomass isocline shifts upward in this region (fig. 3a). This increase is exaggerated in figure 4d because the steady state is near the peak of BI .

Biomass level B_{a1} tracks \underline{B} closely in the four simulations. When \underline{B} is at its low level of approximately 3,000 kilograms per acre (figs. 4a, c), B_{a1} equals 2,974 kilograms per acre (low c_0)

⁷ The location of the upward convergent separatrix associated with (B_H, R_H) , for example, was approximated by fixing $B = B_H$ and repeatedly numerically solving (11) for iterated values of R_0 . The procedure stopped when consecutive recovery levels produced trajectories veering away from steady state in opposite directions. The convergent separatrix is sandwiched between these trajectories. Hence, although the trajectory emanating upward from SS_M in figure 4c appears to be a single trajectory eventually splitting into two branches, it is actually two trajectories beginning from slightly different initial conditions.

⁸ The interval of biomass levels around B_c is the biomass-axis intercepts of the trajectories bounding the convergent separatrix.

or 3,467 kilograms per acre (high c_0). For high $B = 9,000$ kilograms per acre, B_{a1} is 9,280 kilograms per acre (low c_0) or 9,160 kilograms per acre (high c_0).

Conclusion

What do the above results imply about the conditions under which private ownership by a representative areca nut producer preserves the *soppinabetta*? The key parameters determining whether the representative producer selects a continuous recovery strategy over an abandonment strategy when their associated feasible domains do not overlap are B_0 (the initial stock level) and B_c , (the stock initializing the optimal continuous recovery strategy). When $B_0 > B_c$, an optimal continuous recovery strategy leading to ecologically balanced use is overall optimal. In contrast, when $B_0 < B_c$, the producer selects the optimal abandonment strategy if $B_0 \in [B_T, B_{a1}]$, or a suboptimal abandonment strategy otherwise. The simulations, run for a wide range of growth and unit cost curves, produce relatively narrow ranges for B_c of approximately 13,868 kilograms per acre (fig. 4b) to 16,812 kilograms per acre (fig. 4c), and for B_{a1} of 3,467 (fig. 4a) to 9,280 (fig. 4b).

Suppose that the average foliage level reported by Gadgil is the stock that the representative producer, just granted property rights to the *soppinabetta*, initially faces, i.e., $B_0 = 11,336$ kilograms per acre. This initial level falls below the calculated range for B_c and above the calculated range for B_{a1} ; hence, the optimal continuous recovery strategy and the optimal abandonment strategy are both infeasible. Only suboptimal abandonment strategies, producing eventual extinction of the stock, are feasible from this initial stock level.

Privatization will not preserve the *soppinabetta* under these circumstances. However, the government could increase the prospects for preservation by imposing a pre-privatization period of regulated use designed to increase initial biomass levels under privatization beyond the calculated range for B_c . The length and efficacy of such a regulatory period depend on how far

the stock is allowed to drop under the current regime and whether it declines below the minimum viable level, B . Clearly, the government must move fast because Gadgil estimates that current average use will deplete the stock within a decade.

[Received November 1989; final revision received July 1990.]

References

- Albert, A. *College Algebra*, 1st ed. New York: McGraw-Hill Book Co., 1946.
- Arrow, K., and M. Kurtz. *Public Investment, The Rate of Return, and Optimal Fiscal Policy*. Washington DC: Resources for the Future, 1970.
- Clark, C. W. *Mathematical Bioeconomics: The Optimal Management of Renewable Resources*. New York: John Wiley & Sons, 1976.
- Cropper, M. "A Note on the Extinction of Renewable Resources." *J. Environ. Econ. Manage.* 15(1988):64–70.
- Davidson, R., and R. Harris. "Non-Convexities in Continuous-Time Investment Theory." *Rev. Econ. Stud.* 48(1981):235–53.
- Gadgil, M., Hedge, and Bhoja Shetty. "Uttara Kannada: A Case Study in Hill Area Development." *Karnataka State of Environment Report 1985–1986, Government of Karnataka, India*, pp. 155–72. Karnataka State, India, 1986.
- Gadgil, M. "Depleting Renewable Resources: A Case Study From Karnataka Western Ghats." *Indian J. Agr. Econ.* 42(1987):376–87.
- Hadley, G. *Linear Algebra*. Reading MA: Addison-Wesley Publishing Co., 1973.
- Hulgol Service Co-operative Society. *North Kanara Eco-Development Project: A Case Study*, pp. 37. Karnataka, India, 1985.
- Kramer, P., and T. Kozlowski. *Physiology of Trees*. New York: McGraw-Hill Book Co., 1960.
- Lewis, T., and R. Schmalensee. "Nonconvexity and Optimal Exhaustion of Renewable Resources." *Int. Econ. Rev.* 18(1977):535–52.
- Quarterly Economic Report of the Indian Institute of Public Opinion, vol. 31, no. 1, p. 6, Sep.–Dec. 1987.
- Smith, V. "The Economics of Production from Natural Resources." *Amer. Econ. Rev.* 58(1968):409–31.
- Takayama, A. *Mathematical Economics*, 2nd ed. New York: Cambridge University Press, 1985.
- Wargo, P. "Insects Have Defoliated My Tree—Now What's Going To Happen?" *J. Arboriculture* 4(1978):169–75.

Impact Targets versus Discharge Standards in Agricultural Pollution Management

John B. Braden, Robert S. Larson, and Edwin E. Herricks

When attempting to protect fish in streams, sediment or erosion targets are inefficient. Use of a habitat suitability target reveals lower cost abatement measures because it accounts for pesticides as well as soil particles. In Lake Michigan case studies, the lower cost measures involve more crop diversity, less use of no-till, and changes on more acres than the solutions based on sediment discharges or erosion rates.

Key words: environmental policy, fish, nonpoint pollution, optimization, targets.

Protecting water quality has been avowed as a major objective of soil conservation programs (U.S. Department of Agriculture). Yet, controlling soil movement continues to be stressed in practice; an example is the Conservation Reserve Program's emphasis on highly erodible lands. The effectiveness of protecting water quality by stabilizing soil is an open question.

This paper quantifies the economic losses from the use of soil movement rather than water quality criteria for the attainment of an important water quality goal—protection of habitat for high-valued fish species. The estimates are based on a case study of Lake Michigan tributaries. The case study also indicates the differences in farming practices when habitat is protected by controlling soil movement alone versus managing both soil and pesticide pollution.

The analysis extends the methods used previously by Braden et al., Crowder and Young, Heimlich and Ogg, Milon, Park and Sawyer, and Park and Shabman. The common theme of those

studies is the linking of land management economics to off-site consequences. The usual aim has been to compare various policies for meeting specified levels of pollutant loads.

The present study goes beyond previous work by considering predicted impacts on fish habitat. The costs and management implications of actual environmental damages, as well as of emissions and pollutant loads, then can be assessed. These three performance measures conform to the policy targets defined by Nichols—emissions, exposure, and damages; however, his "damage" category implies economic evaluation of the impacts, while no such evaluation is attempted here. Insight into the inefficiencies introduced by inexact targets is important in evaluating whether to incur the additional costs of measuring and monitoring actual damages.

Economic Model

The economic model portrays a fully informed watershed planner. The planner's objective is to achieve environmental goals at least cost. Difficulties of attaching values to environmental impacts frequently lead to the use of such second-best cost-effectiveness criteria (Baumol and Oates).

The decision context involves stochastic risks of environmental impacts. The environmental consequences of interest depend in part on the severity and timing of weather events in relation to crop cover and chemical use. All of these factors can vary over a planning horizon. Under

John B. Braden is a professor, Department of Agricultural Economics, University of Illinois, and a visiting professor, Centre for Resource Management, Lincoln University, New Zealand; Robert S. Larson is a research associate, Illinois State Water Survey; and Edwin E. Herricks is an associate professor, Department of Civil Engineering, University of Illinois.

Support for this research was provided through Project SG-083 of the Illinois-Indiana Sea Grant Program, U.S. Department of the Interior; Cooperative Agreement No. USDA 58-319v-6-00014 of the Resources and Technology Division, Economic Research Service, U.S. Department of Agriculture; and Project 0334 of the Agricultural Experiment Station, College of Agriculture, University of Illinois.

Thanks without implication go from the authors to Roy Black and Wayne Whitman for help with the case studies and to two anonymous referees for suggestions that improved the exposition.

such circumstances, the planner's decision framework follows in spirit the models of Beavis and Walker and Lichtenberg and Zilberman. Let $C(x)$ be a primal cost function defined over choice variables in the vector x , $q(x, \varepsilon)$ be an environmental quality production function, ε be a stochastic disturbance term, Q be a target level of minimum environmental quality, and A be a measure of reliability (the probability of achieving Q). The cost-effectiveness decision is

$$(1) \quad J = \text{Min } C(x), \text{ s.t. } \Pr[q(x, \varepsilon) \leq Q] \geq A.$$

This framework anticipates that higher quality can be achieved only by sacrificing cost or reliability or both. Greater reliability presents similar tradeoffs.

The environmental quality and reliability targets in (1) may be difficult to measure and monitor. Contributing factors, such as emissions or pollutant loads reaching the stream, may provide alternative policy targets. Suppose that the function $q(\bullet)$ can be rewritten as

$$(2) \quad q = \hat{q}(x, h(\hat{x}), \varepsilon), \hat{x} \subset x, x \not\subset \hat{x},$$

where $h(\bullet)$ is an intermediate environmental quality indicator, such as emissions or pollutant loads. The planner then may consider the alternative policy:

$$(3) \quad J = \min C(x), \text{ s.t. } \Pr[h(\hat{x}) \leq H] \leq B,$$

where the intermediate objectives H and B are based on knowledge about the stochastic relationship between $h(\bullet)$ and final objectives Q and A .

For achieving Q and A , the management choices based on (3) will never be less costly than selections based on (1). Fewer choice variables (instruments) are relevant to the performance measures, and those measures are inexact proxies for the final objectives. The extent of the inefficiency and the nature of the management miscalculations are empirical questions.

Model Implementation

The empirical model has been described in detail by Braden, Herricks, and Larson; and Larson, Herricks, and Braden. It consists of a version of the Sediment Economics (SEDEC) model (Braden et al.; Bouzaher, Braden, and Johnson; Johnson et al.) that is extended to include (a) seasonality of sediment loads; (b) pesticide losses, toxicity, and concentrations; and (c) effects of

sediment and pesticides on the habitat requirements of selected fish species.

Briefly, pollutant transport submodels for sediment and pesticides relate farming practices to pollutant delivery to waterbodies. SEDEC places these relationships in a spatial optimization framework, permitting the planner to identify the optimal type and location of interventions to achieve particular goals.

The pollutant loads are translated into habitat quality and reliability through habitat suitability models (U.S. Fish and Wildlife Service). These models are calibrated for individual species, of which more than forty have been characterized. Habitat suitability indices (HSI) are unitless numbers ranging from zero (poor) to one (excellent). They capture the combined effects of relevant habitat parameters, such as temperature, substrate conditions, and concentrations of contaminants.

Timing is extremely important in determining the effects on fish of soil and farm chemicals washed into streams by storm events. Monte Carlo simulations based on historical weather records, planting dates, and chemical application practices are used to capture these stochastic factors. The simulation outputs are probability distributions of pollutant discharges and habitat impacts. The probability distributions identify the likelihood (reliability) of achieving any specific suitability index level with a particular watershed management scenario.

The mathematical expression for the implemented model is a modified and elaborated version of problem (1) above. In addition to the earlier notation, let counting index $j, j = 1, \dots, J$, refer to subwatersheds and index $i, i = 1, \dots, I^j$, denote all possible combinations of management practices on the land units within a particular subwatershed. Each i will be called a management path. Variable $x_{ij} \in [0, 1]$ is binary and represents a management path in subwatershed j . Because only one such path can be chosen in each subwatershed, $\sum_i x_{ij} = 1$ for all j . PC_{ij} is the probability of pesticide concentrations exceeding a particular suitability level Q^* ; PS_{ijg} is the probability of sediment suitability exceeding Q^* in season g ; and A is the target level of reliability. Variable a_j is area of subwatershed j and f_{ij} is the predicted median runoff from management path i of subwatershed j . The decision problem is:

$$(4a) \quad \text{Min } C(x) = \sum_{j=1}^J \sum_{i=1}^{I^j} c_{ij} x_{ij}$$

s.t.

$$(4b) \quad \frac{\sum_{j=1}^J \sum_{i=1}^{I'} PC_{ij} f_{ij} x_{ij}}{\sum_{j=1}^J \sum_{i=1}^{I'} f_{ij} x_{ij}} \geq A$$

$$(4c) \quad \frac{\sum_{j=1}^J \sum_{i=1}^{I'} PS_{ijg} a_{ij} x_{ij}}{\sum_{j=1}^J \sum_{i=1}^{I'} a_{ij} x_{ij}} \geq A \quad g = 1, \dots, G$$

$$(4d) \quad \sum_{i=1}^{I'} x_{ij} = 1 \quad j = 1, \dots, J$$

$$(4e) \quad x_{ij} = [0, 1].$$

Constraints (4b) and (4c) are weighted-average probabilities of exceeding the target suitability level for pesticides and sediment, respectively. Variables a_j and f_{ij} are weighting terms used to reflect the relative contribution of each subwatershed.

A solution to this problem reveals the management choices that will achieve specified habitat quality and reliability levels at least cost. Varying the constraints will show the trade-offs between cost, quality, and reliability. More important, the model can be modified to constrain emissions or pollutant loads. The resulting solutions can be run through the habitat simulations to reveal how close they come to protecting quality and reliability. Doing so reveals the efficiency losses due to the use of intermediate targets.

Different targets should be good substitutes for policy purposes if they are closely correlated (Nichols). In that case, management prescriptions optimal for one indicator should be nearly optimal (although possibly second-best) for the other indicator.

In the case of agricultural pollution and fish habitat, timing can cause critical differences between emissions, releases, and habitat suitability. The timing effects come through decay in pesticide toxicity and the seasonal patterns of rainfall, erosivity, crop conditions, and fish spawning requirements.

Another source of divergence between targets is oversimplification. Environmental damages are frequently the outcome of complex processes involving many factors. Targets based on a subset of the processes may miss some im-

portant contributing factors. For example, erosion rates and sediment loads may not effectively represent the fates of soluble pesticides.

The size of the losses due to the use of proxy targets is an empirical matter. Insights are developed in a case study of sport fish protection in tributaries to Lake Michigan.

Case Study

Active sport fisheries have been successfully developed in Lake Michigan over the past two decades with substantial economic benefits for the near-shore area. Chinook, coho, steelhead, and other salmonids are the most prized varieties.

The salmonid populations have been sustained and enhanced through extensive stocking. Natural spawning has been limited in many tributaries by nonpoint pollution from farmland and by channelization that eliminates habitat while enhancing drainage. These factors not only compel continued stocking, they also reduce the range of seasonal salmon migrations. The migrations are highly valued by individuals and communities near the lake who seek to lengthen and enhance the fish runs.

The model was applied to two agricultural subwatersheds in Berrien County, Michigan, along Lake Michigan's southeastern shore. The Pipestone Creek site drains to the St. Joseph River and on to Lake Michigan. The states of Michigan and Indiana are working together to extend the salmon runs in the St. Joseph River system. The river and its tributaries have been abundantly stocked with juvenile sport fish in recent years. Portions are classified as trout streams. However, the segment of Pipestone Creek chosen for study has been channelized and the silty substrate is poor for spawning and fry development. The 93 hectare (ha) study site contains gently sloping farmland with silty and loamy soils.

The Galien River (east branch) site is part of a smaller river system that also is classified for trout. The habitat conditions are good for salmonids with a meandering channel, cobble and gravel substrate, and pools interspersed with riffle segments. The study site contains 139 hectares of gently sloping farmland with sandy and loamy soils.

Data

Catchments and transects were defined from U.S. Geological Survey topographic maps. Manage-

ment units were identified from Soil Conservation Service (SCS) soil survey maps, plat maps, and Agricultural Stabilization and Conservation Service aerial photographs. Soils information, including productivity classifications, also came from soil surveys (U.S. Soil Conservation Service). Rainfall distributions were based on a fifty-seven-year record for nearby Eau Claire, Michigan. Basic stream data were compiled through fieldwork.

Coefficients for the Universal Soil Loss Equation (Wischmeier and Smith) and crop budgets, including pesticide application rates and assumptions about the timing of farming operations, were prepared by experts in the Michigan Cooperative Extensive Service and the SCS (J. Black, Dep. Agr. Econ., Michigan State University, personal communication 1988). Corn, grains, and soybeans are the most common farm crops in Berrien County, although orchards, vegetables, and vineyards also are present. The crop-cover (*C*) factors for the USLE were disaggregated for crop growth phases, and variability was introduced following Thomas, Snider, and Langdale. Twelve possible cropping systems were considered, consisting of combinations of two rotations—wheat-corn (3)-soybeans (WCCS) and alfalfa (3)-corn (2) (AAACC), three tillage methods—moldboard plowing, till-planting, and no-till, and two mechanical practices—vertical plowing and contour plowing. These options are typical of the area, and the rotations make use of similar pesticides. Three pesticides were selected for study: Atrazine, Furadan, and Bladex. Atrazine and Bladex use does not vary with tillage practices, while Furadan is used in fewer years when tillage is reduced. Assumed crop prices were \$60 per ton for alfalfa hay, \$2.25 per bushel for corn, \$5.40 per bushel for soybeans, and \$2.30 per bushel for wheat.

Chemical toxicity data for salmonids were obtained from Mayer and Ellersieck and incorporated into habitat relationships using the techniques developed by Herricks and Braga. Physical suitability relationships were adapted from existing HSI models (e.g., Raleigh et al.).

Analysis

The SEDEC model was used to determine the economically optimal management practices for meeting sediment targets. The consequences for fish habitat suitability of the practices that optimally control sediment were determined using

the extended model (without optimizing for suitability impacts). A similar approach was followed for erosion targets. The analysis also was performed in reverse—the optimal practices were determined with respect to suitability targets and the sedimentation and erosion consequences of those practices were traced. Finally, the sub-watershed suitability target was applied to all individual catchments to assess the consequences of imposing uniformity throughout the stream reach.

While any suitability, sedimentation, or erosion levels could be selected for analysis, levels of 0.5, 0.7, and 0.9 were chosen here. These cover average to very good suitability conditions, on the assumption that poor conditions are not relevant environmental targets.

Results

The results are summarized as cost frontiers relating the minimum losses in farming profits associated with attaining particular environmental targets. The cost estimates do not reflect differences in farmer risks that may accompany different management systems; they assume that watershed management can be highly selective; and they assume all farmers would settle for the minimum compensation.

Figure 1 shows the cost curves for the two study sites assuming the extreme habitat suitability targets of 0.5 and 0.9 and allowing reliability to vary. The costs are per hectare for comparison, although the costs are borne unevenly across management units as a result of the optimization.

The curves are quite different for the two sites, and this is attributable to the different background conditions. The Galien site is already highly suitable and reliable for salmonids, while Pipestone is not. Thus, the costs are greater for attaining high reliability levels at Pipestone.

The curves for the 0.5 suitability level extend to higher levels of reliability than do the 0.9 curves. This suggests that the best practices for usual weather circumstances (that dominate the suitability determination) are not the same as the best practices for extreme conditions (that dominate reliability). Furthermore, conservative farming practices alone cannot achieve high levels of suitability with high reliability. The dual extremes would require either land use changes more substantial than those considered here or supplementary measures within the stream channels.

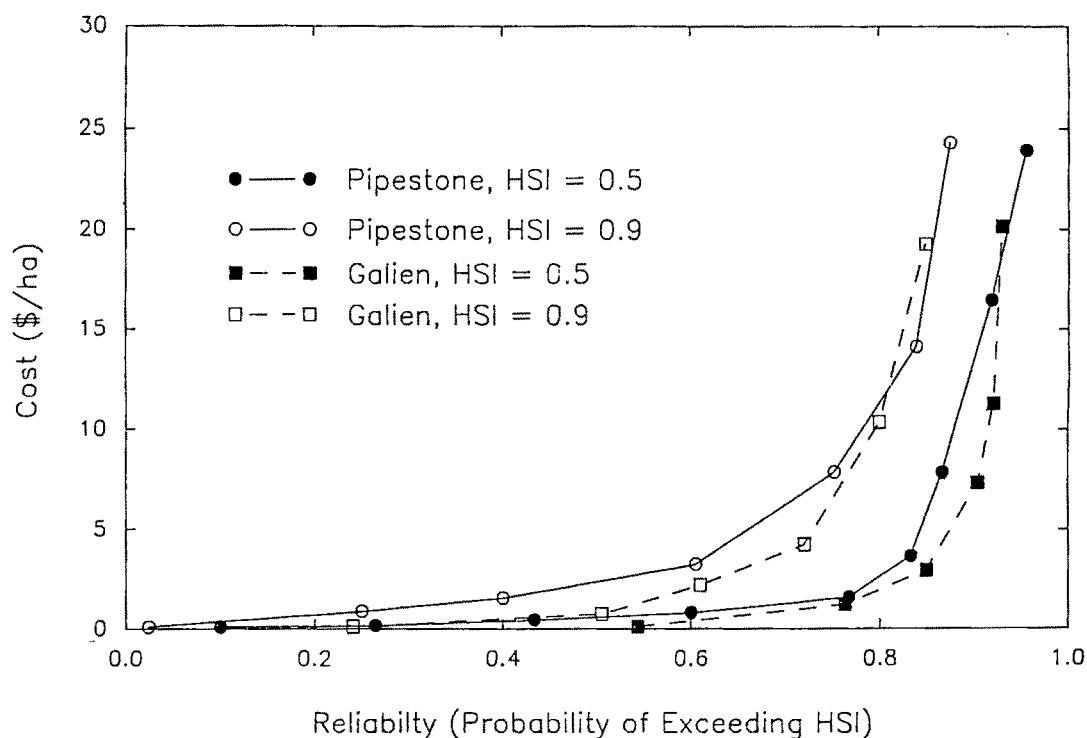


Figure 1. Minimum costs of achieving selected salmonid habitat suitabilities and reliabilities, Pipestone Creek and Galien River sites

The constraint on pesticide suitability is non-binding at low levels of reliability. The pesticide constraint does not become binding until rather high probabilities of exceeding the target suitability levels are reached, at which point the risk of excessive sediment accumulation is relatively low. (The reliability level at which pesticides become important varies inversely with the suitability level). These findings are consistent with the consensus among fisheries biologists that deteriorated substrate conditions are most responsible for the general degradation of fish populations in midwestern streams (e.g., Smith).

For the comparison of targets, figures 2 through 5 display the cost-suitability frontiers for (a) targeting directly on suitability; (b) constraining the total sediment load in the watershed; (c) constraining the sediment load from each catchment; and (d) constraining the soil erosion on each LMU, and the frontier from targeting directly on suitability. The 0.5 and 0.9 levels of suitability illustrate extremes.

The figures show that a sediment target reasonably approximates a habitat suitability target only over a limited range. The approximation grows worse as pesticides play a greater role in suitability determination, i.e., at higher levels of reliability where the pesticide suitability con-

straint is controlling. Because the critical pesticides are in solution, and because sediment runoff is not necessarily correlated with runoff volume or concentration, "targeting" sediment is a poor way to deal with pesticide effects.

Comparisons of the figures suggests that the range of reasonable approximation shrinks as the suitability target is raised. This shrinkage occurs because the pesticide constraints bind at lower reliability levels when suitability targets are higher.

The sediment and erosion target curves in figures 2 through 5 are not smooth because some strategies (e.g., alfalfa rotations) used to control soil movement also lower pesticides while others (e.g., no-till) can increase pesticide concentrations in runoff. Erosion and sedimentation targets take no account of the pesticide consequences and result in higher costs and greater or lesser reliability depending on the nature of the sediment control regime.

Management Implications

Optimal management scenarios for the HSI target are summarized in table 1, and corresponding results for the alternative targets appear in

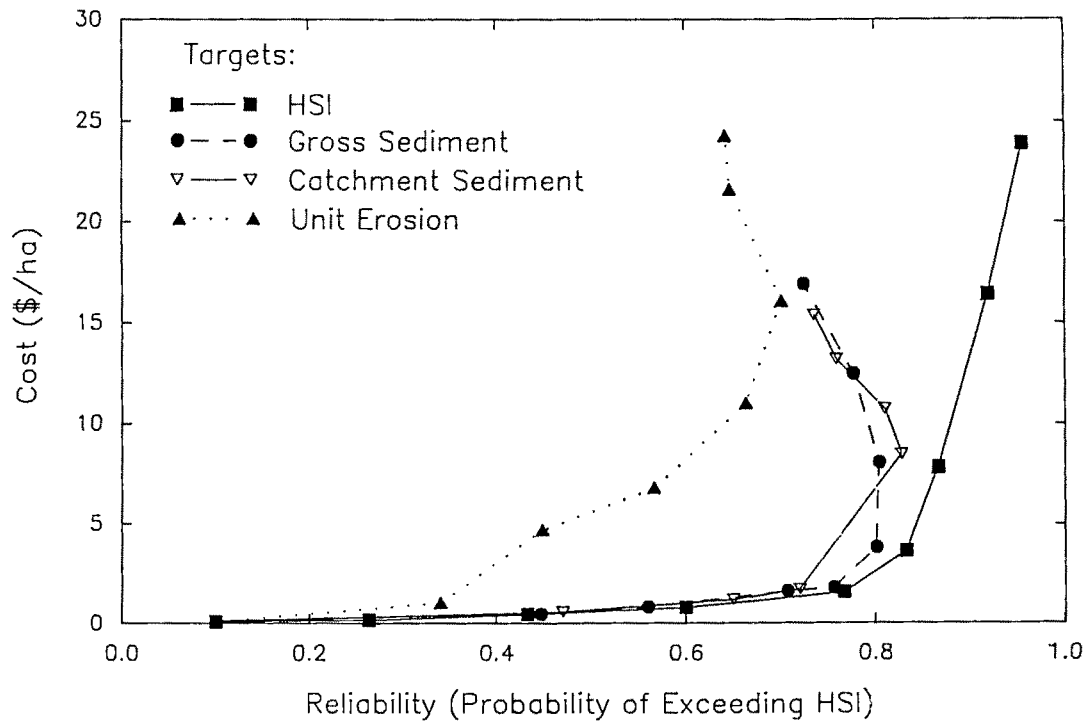


Figure 2. Cost of salmonid reliability with selected discharge targets and impacts, Pipestone Creek, HSI = 0.5

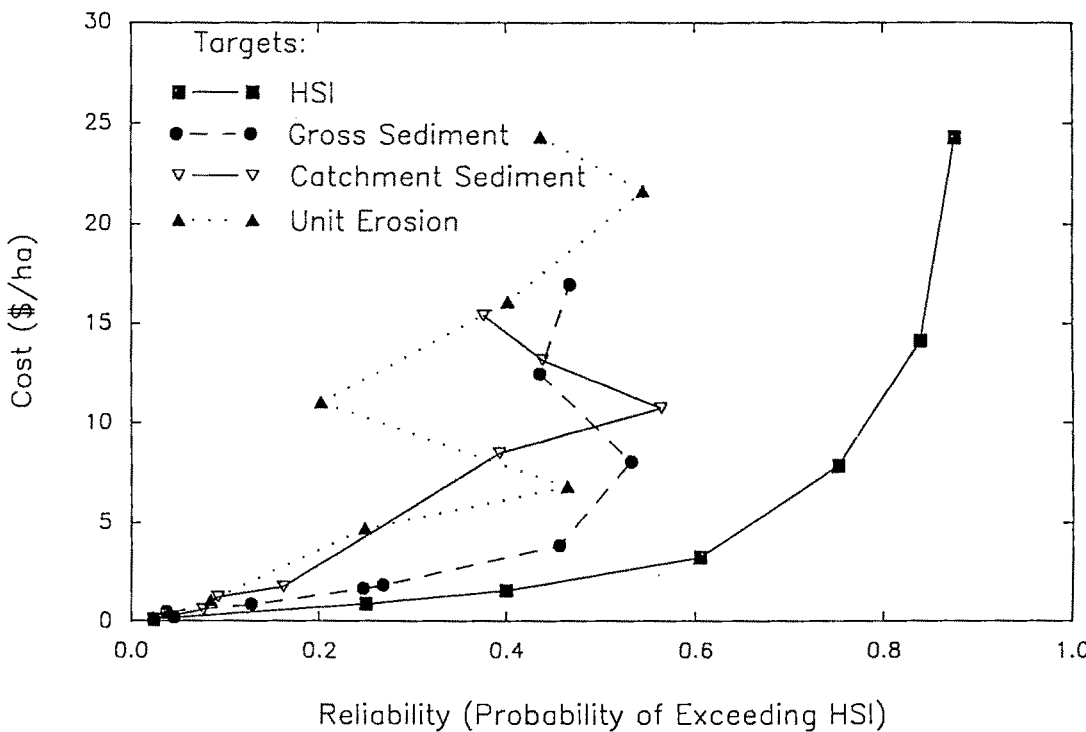


Figure 3. Costs of salmonid reliability with selected discharge targets and impacts, Pipestone Creek, HSI = 0.9.

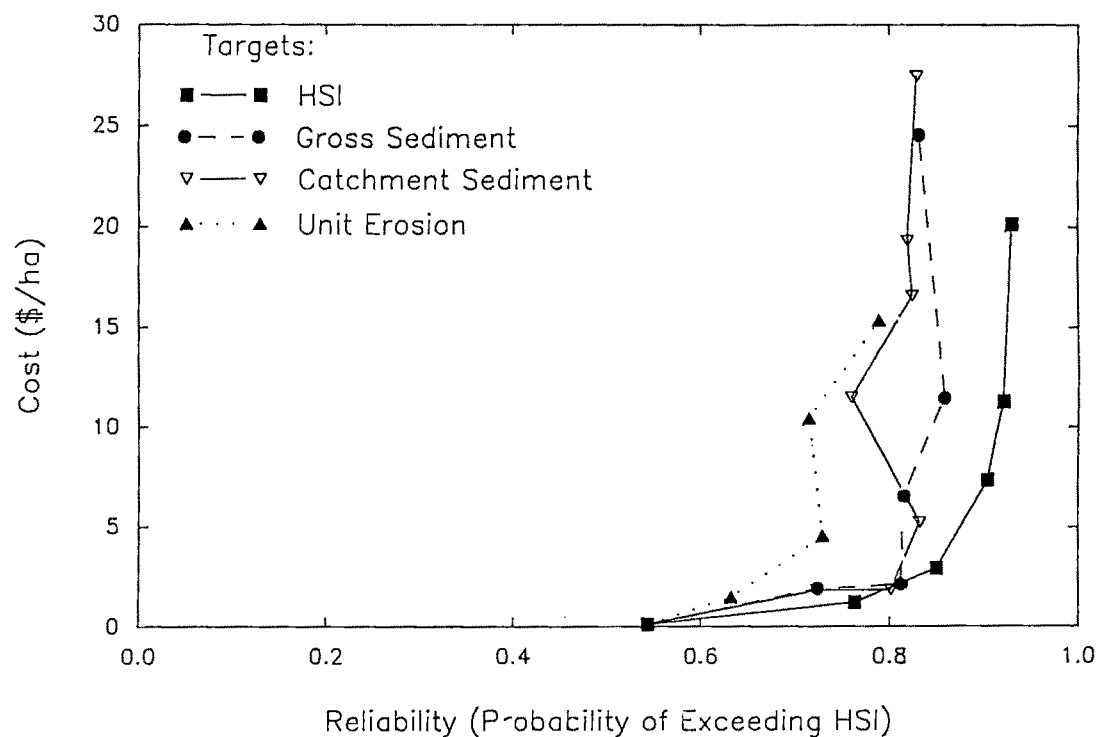


Figure 4. Costs of salmonid reliability with selected discharge targets and impacts, Galien River, HSI = 0.5

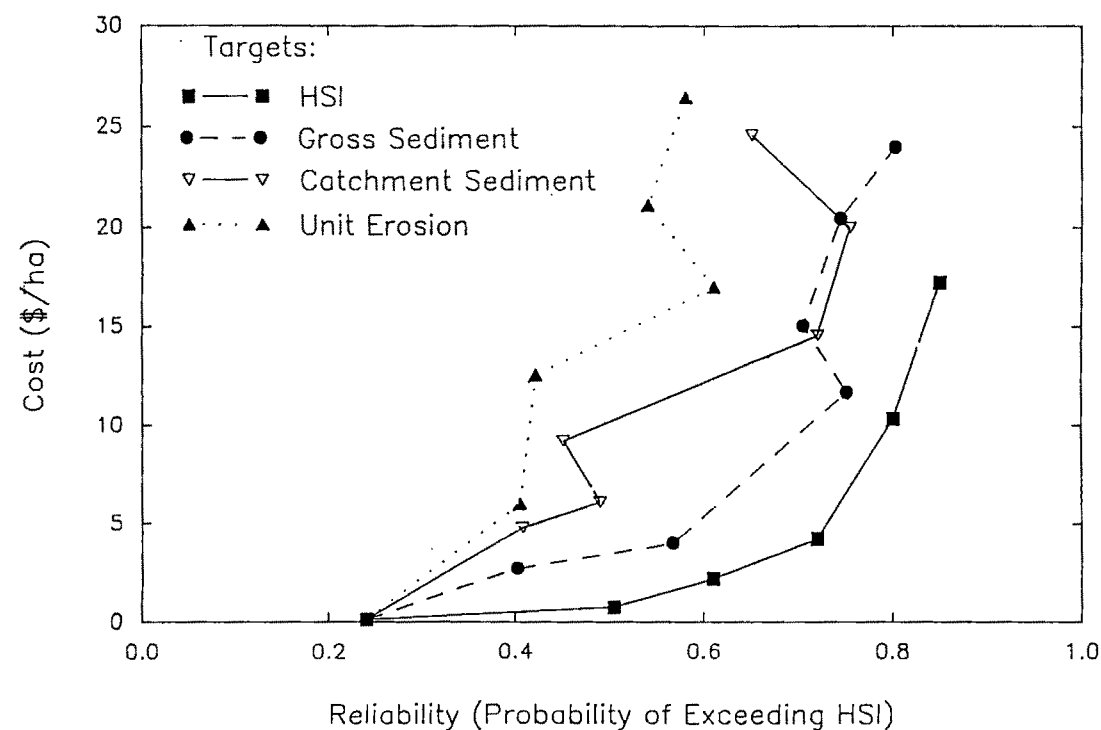


Figure 5. Cost of salmonid reliability with selected discharge targets and impacts, Galien River, HSI = 0.9

Table 1. Optimal Management Summaries for Selected Salmonid Suitability/Reliability Levels, Pipestone Creek and Galien River Sites

Salmonid HSI/Reliability	Management Practices			Extent of Management Changes		Cost
	Rotation	Tillage	Mechanical	(% Area)	(% units)	
	(% ha WCCCS)	(% Mold- board)/ (% No-till)	(% Verticle)			(\$/ha)
Pipestone Creek						
Baseline	100	67/6	93			0.00
0.5/0.40	100	55/17	93	13	5	0.43
0.5/0.80	91	27/28	93	33	11	2.30
0.9/0.40	92	48/28	93	30	26	0.57
0.9/0.80	72	18/31	93	71	80	11.03
Galien River						
Baseline	100	42/22	87			0.00
0.5/0.40	100	42/22	87	0	0	0.00
0.5/0.80	82	26/26	84	41	17	1.60
0.9/0.40	87	35/24	87	33	18	0.45
0.9/0.80	64	24/17	84	65	73	8.77

table 2. The selection of performance goals for table 2 was limited because some or all of the alternative targets could not achieve reliability of 0.8 at Pipestone with either the 0.5 or the 0.9 HSI, nor at Galien with 0.9 HSI.

In the baseline case, without habitat constraints, the WCCCS rotation and a combination

of tillage practices are implemented at both sites. As indicated in table 1, tightening the habitat constraint initially (the 0.5/.40 case) prompts greater use of no-till WCCCS. Requiring reliability of .80 causes a shift away from no-till WCCCS and toward the AACCA rotation. The greater availability and concentration of pesti-

Table 2. Comparison of Watershed Management for Selected Alternative Pollution Abatement Targets and Impacts, Pipestone Creek and Galien River Sites

Impact HSI/Reliability	Target	Management Practices			Extent of Management Changes		Cost
		Rotation	Tillage	Mechanical	(% Area)	(% units)	
		(% ha WCCCS)	(% Mold- board)/ (% No-till)	(% Verticle)			(\$/ha)
Pipestone Creek							
0.5/0.40	Gross Sed.	100	55/17	93	5	2	0.43
	Catch Sed.	100	53/19	93	7	3	0.50
	Unit Erosion	100	51/26	93	13	8	4.39
0.9/0.40	Gross Sed.	92	51/26	93	21	15	2.51
	Catch Sed.	93	53/31	93	24	19	7.87
	Unit Erosion	96	49/44	93	36	25	5.47
Galien River							
0.5/0.80	Gross Sed.	100	42/20	87	2	1	1.73
	Catch Sed.	100	40/20	87	5	2	1.84
	Unit Erosion	100	35/28	87	12	21	2.63
0.9/0.40	Gross Sed.	89	44/32	87	27	8	2.72
	Catch Sed.	90	41/33	87	31	12	4.81
	Unit Erosion	94	33/41	87	37	33	6.02

cides with no-till accounts for this shift. Tightening the constraints also requires that changes be made in more management units and more acres.

For each site, the mechanical practices change little or not at all with different constraints because contour and vertical plowing perform very similarly on the long gentle slopes of the sites.

In comparing tables 1 and 2, the erosion and sediment targets generally lead to more acreage in the WCCCS rotation and more no-till. (An exception to the no-till result appears in the Galien 0.5/.80 case, but more use of conservation tillage with the HSI/Reliability target accounts for this apparent anomaly.) These results are as expected and are more pronounced, respectively, for gross sediment, catchment sediment, and erosion, that is, as the target becomes further removed from fish habitat.

An unexpected result, at least for the sediment targets, is that less area and fewer management units are involved in the solutions, albeit at higher overall costs. An interesting implication is that if administrative costs increase with the area and number of farms involved in abatement actions, the ostensible efficiency gains of using suitability/reliability targets could be offset.

Conclusions

This study suggests that protecting fish habitat can be quite distinct from reducing agricultural erosion or sediment discharges. Policies that address sediment or erosion effectively are less effective in protecting habitat, especially at high suitability and reliability levels. This is because soluble pesticides dominate extreme suitability and reliability conditions, and the correlation to sediment loads is not high. This result is not surprising because fish respond to multiple qualities of the stream channel. Single-dimensional policies will be effective only if the dimension chosen is highly correlated with overall suitability.

A specific policy concern involves no-till farming. No-till has been widely encouraged. At least in the cases studied here, this approach appears sound with respect to erosion and sedimentation. But the consequences for fish, and perhaps other wildlife, may be perverse; no-till sometimes involves greater use of pesticides, which are not as fully incorporated, while it also reduces runoff volume. Nonincorporation means that less water will move more chemicals. The

results point toward the desirability of no-till systems that better control pesticide releases.

Another policy issue involves the apparent desirability of heterogeneous cropping systems in a watershed. When suitability and reliability goals are high, changes in tillage and mechanical practices are inadequate. Crop changes are needed (unless stream channel measures are undertaken), and the changes entail more diversity. More diversity reduces the probability of any one chemical exerting influence in a particular weather event. Agricultural policies favoring the cultivation of fewer crops may hamper efforts to attain high quality stream fisheries in some areas.

The results suggest that less area and fewer farms are affected by targeting on sediment than on suitability. Thus, the apparent disadvantages of sediment targets may be less pronounced when administrative costs are considered.

Finally, the differences between targets in the costs of attaining particular quality/reliability goals are the potential gains from intensive water quality monitoring and measurement to fine tune abatement efforts. Intensive programs will undoubtedly be quite costly. There is apparently little to gain when quality/reliability goals are modest or when existing conditions are generally good. The emissions or exposure targets perform reasonably well. For fish habitat purposes, the gains from intensive programs probably warrant the costs only where they stand a good chance of greatly improving habitat for high-value species.

[Received August 1989; final revision received June 1990.]

References

- Baumol, W., and W. Oates. *The Theory of Environmental Policy*, 2nd ed. Cambridge and New York: Cambridge University Press, 1988.
- Beavis, B., and M. Walker. "Achieving Environmental Standards with Stochastic Discharges." *J. Environ. Econ. Manage.* 10(1983):103-11.
- Bouzaher, A., J. Braden, and G. Johnson. "A Dynamic Programming Approach to a Class of Nonpoint Source Pollution Control Problems." *Manage. Sci.* 36(1990): 1-15.
- Braden, J., A. Bouzaher, G. Johnson, and D. Miltz. "Optimal Spatial Management of Agricultural Pollution." *Amer. J. Agr. Econ.* 71(1989):404-13.
- Braden, J., E. Herricks, and R. Larson. "Economic Targeting of Nonpoint Pollution Abatement for Fish Habitat Protection." *Water Resour. Res.* 25(1989):2399-2405.

- Crowder, B., and C. Young. "Soil Conservation Practices and Water Quality: Is Erosion Control the Answer?" *Water Resour. Bull.* 23(1987):897-902.
- Heimlich, R., and C. Ogg. "Evaluation of Soil-Erosion and Pesticide-Exposure Control Strategies." *J. Environ. Econ. Manage.* 9(1982):279-88.
- Herricks, E., and M. Braga. "Habitat Elements in River Basin Planning." *Water Sci. Technol.* 19(1987):19-29.
- Johnson, G., D. White, A. Bouzaher, and J. Bradén. "SE-DEC User's Guide, Ver. 1.0." Inst. Environ. Stud. tech. rep., University of Illinois, 1989.
- Larson, R., E. Herricks, and J. Bradén. *Efficient Protection of Fisheries Habitat in Great Lakes Tributaries from Agricultural Pollutants*. IL-IN. Sea Grant Program Res. Rep. No. IL-IN-SG-R-90-1, University of Illinois, March 1990.
- Lichtenberg, E., and D. Zilberman. "Efficient Regulation of Health Risks." *Quart. J. Econ.* 103(1988):167-78.
- Mayer, F., and M. Ellersieck. *Manual of Acute Toxicity: Interpretation and Database for 410 Chemicals and 66 Species of Freshwater Animals*. Washington DC: U.S. Fish and Wildlife Serv. Resour. Pub. No. 160, 1984.
- Milon, J. "Optimizing Nonpoint Source Control in Water Quality Regulation." *Water Resour. Bull.* 23(1987):387-96.
- Nichols, A. *Targeting Economic Incentives for Environmental Protection*. Cambridge MA: MIT Press, 1984.
- Park, W., and D. Sawyer. "Targeting Soil Erosion Control Efforts in a Critical Watershed." Washington DC: U.S. Department of Agriculture, Econ. Res. Serv., ERS Staff Rep. No. AGES850801, 1985.
- Park, W., and L. Shabman. "Distributional Constraints on Acceptance of Nonpoint Pollution Controls." *Amer. J. Agri. Econ.* 64(1982):455-62.
- Raleigh, R., T. Hickman, R. Solomon, and P. Nelson. *Habitat Suitability Information: Rainbow Trout*. Washington DC: U.S. Fish and Wildlife Serv. Rep. No. FWS/OBS-82/10.60, 1984.
- Smith, P. *The Fishes of Illinois*. Champaign: University of Illinois Press, 1978.
- Thomas, A., W. Snider, and G. Langdale. "Stochastic Impacts on Farming: V. Risk Adjustment Through Conservation Planning." *Amer. Soc. Agr. Eng. Trans.* 31(1988):1368-73.
- U.S. Department of Agriculture. *The Second RCA Appraisal*. Washington DC, 1989.
- U.S. Department of Agriculture, U.S. Soil Conservation Service. *Soil Survey of Berrien County, Michigan*. Lansing MI, 1980.
- U.S. Fish and Wildlife Service. *Habitat Evaluation Procedures (HEP)*. Div. Ecol. Sci. Rep. ESM 102, Washington DC, 1980.
- Wischmeier, W., and D. Smith. *Predicting Rainfall Erosion Losses: A Guide to Conservation Planning*. Washington DC: U.S. Department of Agriculture Handbook 537, 1978.

Tax Reform and Land-Using Sectors in the U.S. Economy: A General Equilibrium Analysis

Roy Boyd and David H. Newman

A computable general equilibrium model of the U.S. economy is used to assess the effects of the Tax Reform Act of 1986 on land-using sectors (forestry and three classes of agriculture). The model's components include twelve production sectors, six consumer groups, a balanced-budget government sector, and a zero surplus foreign sector. In relative terms, Tax Reform reduces total value added output in land-using sectors to a greater extent than other sectors in the economy. Experiments are also performed comparing partial and general equilibrium specifications of the economy and the impact of choice of input substitution elasticities.

Key words: general equilibrium, land use, tax policy, Tax Reform Act of 1986.

The sweeping changes brought about by the Tax Reform Act of 1986 (TRA '86) directly impacted more economic activities in the United States. Among land-using sectors, numerous tax provisions were changed, and these changes have been the subject for several partial equilibrium (PE) analyses. Income from timber harvests, for example, is no longer eligible for preferential capital gains treatment. Both Guertin and Rideout, and Greene and Kleunder calculated substantial private revenue losses due to higher costs when applying the new (higher) tax rates to a number of different land quality and ownership classes. Similarly, because farmers can no longer benefit from the Investment Tax Credit (ITC) or accelerated depreciation, producer profit-maximization models by Leblanc and Hrubovcak, and Boehlje and Carman predicted much higher capital costs for various crops in the future. A fundamental shortcoming of these studies is that they focus on a particular sector, separate from the rest of the economy. The effects of tax changes are economy-wide, and the resulting changes in factor and output prices can have profound effects on individual sectors, apart from their direct tax consequences.

Computable general equilibrium (CGE) models simulate the direct and indirect effects of proposed government tax legislation on a variety of industries and consumer income classes. Unfortunately, these models generally disaggregate manufacturing sectors while aggregating those sectors (such as agriculture and forestry) involved in the production of raw materials. Thus, understanding of the consequences of specific policies on land-based sectors is reduced. Our model closely follows Ballard et al.'s (BFSW) large CGE model of the U.S. economy in that their methods are used for constructing inputs into the model; however, our model is smaller and more specifically related to the policy concerns of this analysis. On the supply side of the economy, total U.S. production is divided into twelve production sectors which are used as inputs to create thirteen consumption commodities (as opposed to 19 and 15 in BFSW).

The purpose of this paper is to quantify the direct and indirect effects of the TRA '86 on land-using sectors using this CGE model of the U.S. economy. The analysis focuses on two important methodological and policy concerns. First, the results are compared using the same model inputs for both CGE and PE analyses. In this way, potential biases derived solely from the methodological approach can be discerned. Related to this analysis, the impact of substitution elasticities assumptions on these results are also examined. As a policy concern, we examine the equity effects of TRA '86 and its im-

The authors are assistant professors, Department of Economics, Ohio University, Athens, and the School of Forest Resources, University of Georgia, respectively.

The authors wish to thank Charles Ballard, Don Fullerton, Thomas Hertel, James Hrubovcak, W. David Klemperer, and Michael LeBlanc for their help and suggestions during various stages of this research.

impact on land-holding classes. The CGE model stratifies consumers by income class, allowing a more accurate portrayal of the direct equity impacts of major changes in the marginal rate structure, as well as the indirect equity effects of changes in relative consumer and factor prices (particularly that of land).

The theoretical, empirical, and tax components of the CGE model are discussed in the next section. Then the results of the policy runs are introduced, focusing primarily on the output, input, and price effects in the land-based sectors. From these results, we discuss the differences derived from equivalent PE runs as well as the sensitivity tests for substitution elasticities. The paper concludes with a discussion of the implications of TRA '86 for land-based sectors in the economy.

Specification of the CGE Model

Land-based production is divided into four sectors: forestry (i.e., timber crops used for further manufacturing production); AG-I, program crops (e.g., corn, wheat, cotton, etc.); AG-II, livestock (e.g., cattle and hogs, etc.); and AG-III, all other agricultural production (e.g., vegetables and fruit crops). Initially, competition for productive factors and simultaneous consumption demand shifts are ignored so that direct, tax induced, changes in these sectors can be measured. This allows the comparison of the PE results with previous work. Integrating these four sectors into the complete CGE model indicates the importance of general equilibrium considerations for land-based production.

Although the concepts involved in CGE models are straightforward, the model itself is fairly complex. In order to facilitate the model's presentation, we use the diagrams in figure 1, the equations in table 1, and the variable acronyms listed in the appendix. The overall equilibrium condition for each production sector j is

$$(1) \quad Y_j + GE_j + VM_j = \sum_j RAS_{jL} + GD_j + VX_j + INV_j.$$

Each of these components is now examined.

Supply

In figure 1, total sectoral output, Y_j , is the product of either a double- or triple-nested production function. First, material inputs enter as

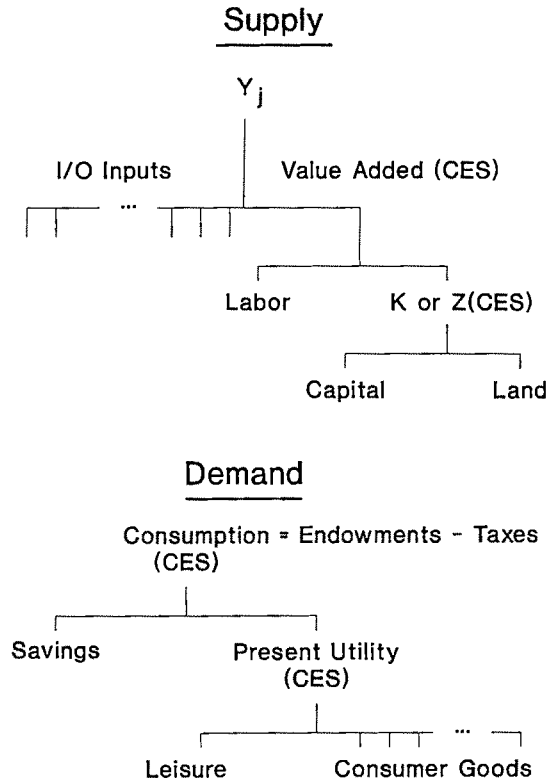


Figure 1. Schematic representation of supply and demand for the empirical model

Leontief, fixed-factor components of the national input-output matrix. The values added to production by labor and capital then enter as primary inputs into a nested constant elasticity of substitution (CES) production function. The substitution elasticity between these inputs can be set to any value between 0 and ∞ .

For the four sectors explicitly using land as an input, a nested variant of the value-added CES function is used. Input Z , which is a CES function of land and capital inputs, replaces K and, in effect, gives a CES production function within a CES production function. Nesting these functions, rather than using a simple three-factor CES value-added function, allows substitution elasticities to vary between the three primary inputs. Within such a setup, PE treatment of these four sectors is straightforward as primary inputs are provided at the (tax-inclusive) price and output demand does not shift. However, under general equilibrium, input prices may vary in response to economy-wide shifts and demand may shift in response to consumer income changes.

The use of land differs from previous CGE

Table 1. Equations for Supply, Demand, and Consistency Aspects of the Empirical Model

Input Equations	
(2) $\sum_c SL_c = \sum_j DL_j + GDL$	(5) $GDL = \sum_j TL_j$
(3) $\sum_c SK_c = \sum_j DK_j + GDK$	(6) $GDK = \sum_j TK_j$
(4) $\sum_c SD_c = \sum_j DD_j + GDD$	(7) $GDD = \sum_j TD_j$
where	
Transfer to Consumer Goods	
(8) $CD_j = \sum_i \Psi_{ji} [GCE_i - TC_i]$	(9) $\sum_c RCS_{ic} = GCE_i$
(10) $\sum_i RCS_{ic} = SL_c + SK_c + SD_c + TRN_c - PIT_c$	
Demand—Household	
(11) $GC_c = \sum_i RCS_{ic} - SAV_c + (1 - \tau_c)(ZTA)SL_c$	
(12) $GC_c = SL_c + SK_c + SD_c + TRN_c - PIT_c + (1 - \tau_c)(ZTA)SL_c$	
Government Tax Adjustment	
(13) $TE = \sum_c [SL_c ZTA_c \tau_c + SK_c \tau_c + SD_c \tau_c - (\Omega_c + TRN)]$ where $\Omega_c = SL_c \tau_c + SK_c \tau_c + SD_c \tau_c - PIT_c$	
Balance—Foreign	
(14) $\sum_j [VM_j \Psi EM_j / (1 + EM_j) + VM_j / (1 + EM_j)] = \sum_j (VX_j + FE_j)$	
Consistency	
Household income equals expenditures	
(15) $\sum_c (SL_c + SK_c + SD_c + TRN_c - PIT_c) = \sum_c (CD_c + TC_c)$	
Government plus income equals outlays	
(16) $\sum_j (GSK_j + GE_j + TL_j + TK_j + TD_j + TXO_j) - \sum_i TRN_i = \sum_j (GDK_j + GD_j) + GDD + GDL$	
Total imports equals total exports	
(17) $\sum_j (VM_j - VX_j) = 0$	
Value of market excess demand equals value added plus taxes	
(18) $\sum_j (CD_j + GD_j + VX_j - CE_j + VM_j) = \sum_j (DL_j + DK_j + TD_j + DD_j + TL_j + TK_j + TXO_j)$	

Note: See appendix for description of variables.

models which included land. Hertel and Tsigas (1987, 1988) subdivided agriculture into seven crops in order to explain specific agricultural subsidies. Kokoski and Smith also used land as an input and were also interested in partial versus general equilibrium estimates. However, their model contains only a single agriculture sector. BFSW performed counterfactual experiments, such as consumption versus income taxes. They felt agriculture played a relatively minor role in such experiments and aggregated it into one sector, subsuming land into a broadly defined capital input.

Primary factor inputs are drawn from each of the ($c = 6$) income groups. Each income group supplies different quantities of land, labor, and capital (SD_c , SL_c , and SK_c), allowing the calculation of gains and losses accruing to specific resource owners. The total of these resources is either consumed as sectoral inputs or taxed away by the government [equations (2)–(7)]. Following BFSW, a unit of an input factor is that quantity which would produce \$1 worth of output per year, net of taxes. This accounting method does not allow the determination of input quantities by output sector, but it does allow the transfer of production units between sectors and deter-

mination of the resulting change in factor shares following a tax change.

Demand—Household

Each household's consumption is modeled as a triple-nested consumption function. At the first level, the consumer faces a CES tradeoff between present and future consumption (savings). Savings financed investment is assumed to equal total savings throughout the economy. Because this analysis is a static simulation, the savings/expenditure choice is not needed to account for dynamic economic growth. Rather, it is included because those taxes which discourage savings and investment have important consequences for output and relative prices in the land-based sectors.

At the second level, the consumer trade-off between leisure time and the goods and services bought from labor services is modeled with a CES nest. Finally, the ($i = 13$) goods and services, including savings, consumed by each household are derived from the ($j = 12$) sectoral outputs and taxed by the government in the form of excise taxes TC_i [equations (8)–(9)]. The

elasticity of substitution between any pair of these goods is assumed equal and, for most cases, is set to 1. Thus, the function collapses into a simple Cobb-Douglas utility function for each household, which is maximized subject to an income constraint [equations (9) and (15)].¹

Demand—Government

Given its initial endowments of goods as well as the taxes it collects, government demands goods and services to the extent that no surplus or deficit remains [equation (16)].² Government purchases should reflect some sensitivity to prices, so government demand for sectoral outputs, capital, and labor is modeled as a single-tiered, Cobb-Douglas consumption function. To preserve the correct incentive structure, consumers in each income class are taxed at their marginal tax rate, τ_c , and then receive the difference between collections using this rate and their average tax rate, PIT_c , as a lump-sum transfer, equation (12).

Foreign Trade

We assume that U.S. and foreign countries either cannot borrow from each other or that such borrowing is constant. Balance of trade thus remains constant [equation (17)] and the import and export elasticities of individual goods then may be varied by adjusting EM_j and FE_j [equation (14)]. Estimates for these elasticities were primarily taken from Stern, Francis, and Schumacher who indicate that "it may be reasonable to work with 'typical' demand elasticities that vary between . . . -0.5 and -2 for exports" (p. 14). We use elasticities close to their estimate of -.85 for agriculture, food, and raw materials such as timber and between -1.24 and -1.41 for other sectors. Finally, if equation (18) holds as an equality, the model is completely specified.

¹ The exact derived demand elasticity from this functional form is given in BFSW. Using their derivation, the share parameters used in this study and aggregating over all income classes, we obtain a derived demand elasticity for all agricultural goods of approximately -0.7. This elasticity is slightly lower than the -0.75 recently derived by Tarr and is the same as earlier estimates by Houthakker. Lower demand elasticities for individual agricultural commodities are presented in Hertel and Tsigas (1988).

² This assumption can easily be relaxed but would require an appropriate change on the consumption-investment equality or the balance of trade constraint.

Base Year Data

The methodology for manipulating the data for the empirical model is explained in BFSW. Data for the twelve production functions come from Caddy, Tarr, and BFSW. Capital receipts and taxes are computed from data provided by the U.S. Department of Commerce (1986), the U.S. Department of Agriculture (1985), Commerce Clearing House, the U.S. Department of Energy (1986a, b) and information provided in Hertel and Tsigas (1987). Factor shares for capital and labor in 1984 were obtained from recent spreadsheets by the Bureau of Economic Analysis. For the land-using sectors, capital was separated from labor by using the Forest Service's (preliminary) 1982 input-output matrix. Land was then separated from capital using spreadsheets from the ERS and the Forest Service.³ Because it would be difficult to derive substitution elasticities for this entire system of equations, these values are obtained from the existing literature. In order to get the sensitivity of the analysis to their values, the simulations were later run with altered elasticity values.

Corporate capital income receives double taxation; it is first taxed as corporate profits and the remaining distributed dividends then are taxed a second time through personal taxes. In order to calculate the total effective tax on capital, we used the method outlined in BFSW (reported in Boyd, table 2). Gross capital income for each industry was first derived, subtracting off property, franchise, and corporate income taxes. The fraction of capital income coming to individuals was then calculated along with the weighted average of the six income classes, τ . These values are then multiplied by capital earnings in the i th industry. Finally, each sector's tax rate is calculated, dividing net capital income by direct plus personal taxes in each sector. These values are shown in table 2. For highly incorporated sectors, such as manufacturing or food processing, double-taxation is more strongly felt. Less incorporated sectors, such as the agriculture and forestry sectors, generally escape the double tax.

In most cases, labor income was computed directly from the salaries and benefits for each production sector. For self-employed persons

³ The raw factor shares (table 3) were similar across all agricultural sectors. A counter intuitive finding was that field crops had the highest labor factor share. This anomaly may be explained by the high wage value of labor's marginal product in that sector and by the large amount of unreported earnings elsewhere in agriculture.

Table 2. Changes in Tax Rates, Value-Added Output, and Net Relative Sector Prices as a Result of Tax Reform Part-1 and Part-2 (in billions of 1984 \$)

Sector	Tax Rates on Capital Inputs		Output Changes				Price Changes ^a		
	BASE	TAX-2	BASE	TAX-1	TAX-2	CHANGE ^b	TAX-1	TAX-2	CHANGE ^b
	----- (%) -----		----- (\$) -----				----- (%) -----		
Manufacturing	73	83	1,887.97	1,886.77	1,886.77	-0.1	1.000	1.000	
Mining	48	50	46.24	46.25	46.28	0.1	1.000	0.991	-0.9
Service	44	51	2,378.08	2,379.99	2,370.99	-0.3	1.000	0.992	-0.8
Chemicals	63	72	227.37	227.47	226.80	-0.3	1.000	0.999	-0.1
Food & tobacco	85	94	350.26	350.91	349.27	-0.3	1.001	0.998	-0.2
Refining	32	49	161.28	161.45	161.66	0.2	1.000	0.991	-0.9
Financial	67	68	554.94	555.73	559.15	0.8	1.000	0.983	-1.7
Crude oil	74	76	129.10	129.22	129.47	0.3	1.000	0.978	-2.2
Forestry	42	42	10.54	10.49	10.39	-1.4	1.004	1.018	1.8
AG-I	31	48	45.09	45.05	44.92	-0.4	1.003	1.002	0.2
AG-II	39	45	109.79	109.94	109.55	-0.2	1.003	1.004	0.4
AG-III	39	41	60.99	60.96	61.21	0.4	1.003	0.991	-0.9
			5,961.65	5,964.23	5,956.46	-0.1			

^a Base case relative prices are set to 1.000 for all sectors.

^b Change is the percentage change from the base case to the value after TAX-2.

(such as sole proprietor or farmer), labor income was imputed by taking the hourly wages of part-time employees in the same field and assigning them to the self-proprietors on a full-time basis (any excess was assigned to land or capital). Labor tax rates were obtained by subtracting from gross labor income, all employer contributions to social insurance and employee contributions to Social Security, unemployment insurance, public workman's compensation, and railroad retirement and dividing this value by net labor income. These rates range from 0.133 and 0.158 in manufacturing and food processing to

0.094, 0.094, 0.098, and 0.105 in the agricultural and forestry sectors. The agricultural rates are lower because many nonincorporated operations do not make social security and other labor payments (Boyd, table 3).

The land-using sectors' land taxes were computed by taking factor shares given by USDA, Economic Research Service (ERS) on capital use and extracting the land component from them. Land tax rates differ from capital tax rates in each sector by the relative magnitude of the subsidies and consumption allowances received by the two factors. The capital tax rates in the 3

Table 3. Marginal Tax Rates, Total Net Income, Relative Price Changes, and Absolute Factor Losses by Income Classes and the Government Sector in the Base Case and New Values Arising from Tax Reform Part-1 and Part-2

Income Class	Marginal Tax Rates		Billions of 1984 \$				Price Index ^b	"Utility" Changes ^c	Million 1984 \$		
	BASE	TAX-2	BASE	TAX-1	TAX-2	Change ^a			Land	Labor	Capital
	----- (%) -----							(%)			
0-9,999	14	3.78	223.83	234.86	233.52	4.3	0.996658	4.0	-11	-408	-1,673
10,000-14,999	16	15.00	210.80	213.56	212.46	0.8	0.996684	0.5	-10	-507	-1,660
15,000-19,999	18	15.00	242.42	248.01	247.03	1.9	0.996726	1.6	-12	-637	-1,648
20,000-29,999	22	20.57	601.31	608.69	607.23	1.0	0.996787	0.7	-28	-1,404	-3,086
30,000-39,999	28	21.85	549.73	564.70	563.85	2.6	0.996841	2.3	-25	-1,215	-2,302
40,000 +	38	28.00	1,373.63	1,409.96	1,402.75	2.1	0.996842	1.8	-41	-4,875	-14,321
Government	—	—	745.79	741.88	758.47	1.7	—	—	—	—	—
Sum			3,947.51	4,021.66	4,025.29	2.0			-127	-7,046	-24,690

^a Change is the percentage change from the base case to the value after TAX-2.

^b Price index is a Laspeyres price index based off of the pre-tax quantities.

^c "Utility" change quantify the combined effects of changes in purchasing power, leisure, and the cost of living.

agricultural sectors and forestry are 0.309, 0.372, 0.395, and 0.418, respectively (Boyd, table 1).

Other data needed came from the Bureau of Labor Statistics, which gave average expenditures for commodity consumption, factor earnings, and taxes paid by income class. The Survey of Current Business of the Department of Commerce (1986b) gave investment/saving, import/export, and government purchase data by sector. Finally, a 12×13 transfer matrix, based on the Department of Commerce's 1977 input-output table of the U.S. economy, described the transformation of sector outputs into commodity goods.

The Simulation Experiment

Initially, the PE simulation experiment is conducted for each of the four land-using sectors. Using the above data, output prices and production levels are solved within the context of a cost-minimization problem. The nested CES production functions are used to solve for these variables (Boyd, p. 7) under constant input prices and output demand. The appropriate input (or output) taxes are then added and the results are compared.

The general equilibrium simulation is more complex and was computed using Rutherford's solution algorithm. As a check of the model's consistency for the 1984 base year, we used a method analogous to that in BFSW's model. A major difference from their model is the inclusion of land as a separate factor of production in our model. As a final check, the base case model was run to determine if the simulation output matches the actual 1984 data. If it does, this solution becomes the benchmark for comparison with the output obtained after the model is run with tax changes.⁴

The policy analysis compares the simulation for the pre-TRA '86 base year of 1984 with the tax structure after the implementation of TRA '86. The taxes are added in two stages. First, only income tax changes affecting consumers (TAX-1) are added. These include changes to personal income and labor taxes (social security) which increase net wealth in all classes (though not by the same amount or percentage)

by lowering the marginal rate. These changes act similarly to a demand shifter for final consumption.

The greatest tax break from TRA '86 change goes to the upper and lower classes. The before and after marginal tax rates faced by the different consumer classes are shown in the first two columns of table 3. The U.S. tax system is only piecewise linear so that the marginal rates in table 3 apply only to the incremental tax dollars paid and not to the average rate paid by each income class (which is generally much lower). In the model, each income group is initially taxed at the relevant marginal rate and funds needed to keep average rates at their 1984 level are returned to them as lump sum payments.⁵

In the second stage (TAX-2), the changes used in TAX-1 are expanded by incorporating the reforms involving input taxes. TRA '86 removed corporate capital gains taxes and various capital recovery (depreciation) loopholes, as well as a general repeal of the investment tax credit (ITC). For most sectors, these changes are incorporated by using calculations from Fullerton, Gillett, and Mackie, as well as by taking values directly from Internal Revenue Service (IRS) tables and adding them into our calculation of the input tax for capital (table 2). For the agriculture sectors we used ERS spreadsheets, based on the IRS breakdowns. As shown in table 2, AG-I and AG-II are affected much more by the loss of the ITC and other tax breaks than is AG-III.⁶ These rate changes are in general agreement with previous work. Hansen and Bertelsen predicted capital price rises ranging from 15% to 32% for program crops and approximately 15% for livestock. No corresponding numbers have been calculated for the AG-III sector.

The differences between the agriculture sectors occur primarily because field crops and livestock require relatively large investments in buildings, storage facilities and heavy equipment. AG-III, on the other hand, produces largely nonstorable crops and relies more heavily on low cost labor for harvest and transport. Forestry was exempted from the provisions removing the ITC, but TRA '86 removed the special capital gains

⁵ The lump sum payments necessary to keep average taxes at their 1984 levels for the post-TRA '86 case are (in billions of dollars) for the lowest to the highest income classes, \$15.4, \$2.0, \$9.5, \$9.8, \$42.4, and \$193.1.

⁶ The formula for including tax changes for nonincorporated capital comes from BFSW (p. 67). The noncorporate ITC lowers the share of each industry's capital payments. The corporate tax is then allocated to each sector by its share and the ITCs used in each sector.

⁴ Although TRA '86 was explicitly designed to be revenue neutral, no revenue-neutral feature was built into our model. Instead, we use an unconstrained simulation because it is more realistic and it conveys valuable information on possible changes in the government deficit.

treatment status.⁷ Hence, in our model, we employ an output tax of 3.38% of the previous output value.

Results

Output Effects

Table 2 shows the expected value-added changes in the twelve production sectors from TAX-1 and TAX-2. While these changes appear small relative to sector size, the actual dollar value changes are substantial. As expected, output in most sectors increases from TAX-1 because of the increase in consumers net income and the resulting outward shift in demand. Only the manufacturing sector and several of the land-based sectors show output declines. Much of the manufacturing sector's output is used for investment, which declines as the relative return to savings falls as a result of the tax change.⁸ This result conforms with the findings of Fullerton, Gillett, and Mackie, who, with a CGE model with a fifty-year time horizon, concluded that the savings loss combined with reduced capital tax incentives chokes off manufacturing investment.

Output in land-using sectors declines because relative increases in the prices of more intensively used factors (capital and land) are greater than the increase in demand for their products. Prices, net of tax, for capital and land increase relative to labor because higher net wages induce more labor supply, making it cheaper for firms relative to the other factors (net, post-TRA '86, factor prices are, 0.9987 for labor, 0.9949 for capital, and 0.9557 for land while the base case rates for all factors are 1.0000).⁹ The biggest gainer from the large drop in the price of capital is the financial sector. As a result, consumption of commodities such as housing (which

strongly affects the forest sector) are expected to increase, because they are dependent on financial service costs.

Government revenue rises while overall output declines with the implementation of TAX-2. The forest sector, as well as several other sectors show relative declines as the loss of subsidies selectively impacts sectors. AG-I and AG-II are particularly hard hit by the loss of ITCs and other tax privileges as indicated by a loss of about \$127 million to landowners (table 3). High income groups bear the highest absolute losses, but simple calculations show that higher proportional losses accrue to lower and middle income landowners. The declines in AG-I and AG-II contradict Fullerton, Gillett, and Mackie, even though our CGE methodology is similar. The contradiction results from the use of different data sources. This study used ERS data which indicate that farmers have heavily utilized the ITC and various other tax benefits. This usage may be masked by more aggregate data sources.

The 1.4% decline in forest sector output is larger than any other sector, but the drop is much smaller than predicted by previous PE analyses (10% and higher). The previous studies did not allow for an overall decrease in the cost of land resulting from its general productivity decline. The loss of capital gains tax preferences leads to an annual increase of \$367 million (1984 \$) in taxes paid, which is the major cause of the drop in sectoral output. This loss is offset to some extent by the retention of ITCs and changes in factor input usage. Previous studies have ignored the change in factor prices and usage in calculating their much larger sector impacts.

Factor usage is set by the gross factor return faced by firms. Even though net capital price declines, the gross price faced by firms buying capital inputs rises relative to labor.¹⁰ Because forestry is relatively capital intensive (table 4), its costs increase relative to other, more labor-intensive sectors like manufacturing. Forestry firms thus shift to more labor-intensive management, rather than capital and land-using management patterns. In the process, they give up 0.39% of their land holdings and experience relatively larger declines in their capital and land factor shares compared with the decline in their labor factor share. This result, though consistent qualitatively with previous studies, is much smaller in magnitude than the large declines in

⁷ In most instances, a capital gain refers to the change in the price of an asset and preferential tax treatment would be modeled as a lower input tax. In forestry, preferential treatment was given to the harvested timber value. Because such gain reflects the value added by the production process itself, the removal of capital gains requires it to be modeled as an output tax.

⁸ Following Boskin, a savings elasticity of 0.4 was used for the empirical simulations. Because of conflicting evidence from Summers and Owen, values of 0.3 and 0.8 were used in the sensitivity analysis. Raising (lowering) the savings elasticity to 0.8 (0.3) caused inconsequential increases (decreases) in agricultural production and prices.

⁹ An aggregate leisure demand elasticity of -0.3 (Killingsworth) was used. In sensitivity tests, this elasticity was lowered to -0.5, which caused the supply of labor to increase slightly, causing the relative prices of land and agricultural goods to rise. These small changes point to the robustness of the model.

¹⁰ Changes in the gross price of capital facing firms can be calculated by adding the changes in net capital price to the changes in factor tax rates given in table. 3.

Table 4. Factor Shares in the Land-Using Sectors Occurring as a Result of Tax Reform Part-1 and Part-2

Factor	BASE*	Forestry TAX-1	TAX-2
		(%)	
Capital	34.2	34.4	33.3
Labor	15.2	15.2	15.1
Land	24.9	25.1	24.3
		AG-I	
Capital	18.8	18.9	19.3
Labor	13.7	13.7	13.7
Land	19.8	20.1	19.8
		AG-II	
Capital	13.5	13.5	13.4
Labor	10.1	10.1	10.1
Land	14.1	14.3	14.1
		AG-III	
Capital	17.1	17.1	16.7
Labor	12.0	11.9	11.9
Land	18.0	18.2	17.8

* Percentages sum to less than 100% because inputs from other sectors are included in the production function.

the productive forest land base predicted by the PE analyses.

Output declines also occur in AG-I and AG-II, which face increased tax payments of about \$640 million. The loss of ITCs is especially damaging because both sectors previously used them extensively. Productive land holdings increase by 0.08% and 0.007% as firms substitute land for capital. Surprisingly, AG-I actually increases its factor share for capital because the relative increase in its gross price is large enough to cause an increase in its total factor share, despite a decrease in the quantity of capital. AG-III, which prior to TRA '86 did not significantly use ITCs, is the only land-using sector showing positive gains in value added production with capital usage increasing and land usage decreasing.

Price Effects

TRA '86's effect on received prices, net of taxes, for sector outputs is consistent with the previous results. The *numéraire* for all prices in table 2 is the manufacturing sector price. Most sector prices drop relative to manufacturing because capital input costs in other sectors are less severely affected by the loss of ITCs and preferential capital gains tax rates.

Prices of forest sector products increase substantially relative to manufacturing. Increases in

the forest sector's operating costs and declines in output combine to increase the price of forest sector products. As a result, demand for imports from the foreign forest sector increases nearly 2%, from \$4.38 billion (1984 \$) to \$4.46 billion (1984 \$). Similarly, the two agricultural sectors exhibiting declining output also show rising prices. Because the overall cost of capital drops, the increases in agricultural prices are smaller than those found by Hansen and Bertelsen.

Partial Equilibrium Simulations

Because PE analyses assume that factors of production are infinitely elastic, overall factor price changes resulting from taxation are not allowed. In PE analyses capital price is fixed and known, while in a CGE model it is part of the solution. Thus, while some of the incidence of an increase in the capital tax falls on capital owners in a CGE model, all factor price increases are shifted forward in a PE analysis. For example, the PE analysis for the forest sector predicts a 3.4% increase in price and 3.1% decrease in output, while the CGE model predicts only a 1.7% increase in price and 1.4% decline output. Similar, yet smaller changes also occur for agriculture (compare the base case and PE columns in table 5).

The CGE and PE results differ for two reasons. One is the microeconomic effect that greater price increases from the supply shift lead to greater quantity declines along the downward-sloping demand curve. Second, a tax change affects demand as well as supply in a CGE model because all markets are connected. In this analysis, the demand for land-based goods shifts out from TRA '86 and further reduces the predicted drop in output.

Sensitivity Tests for Substitution Elasticities

A major concern in studies such as this is the need to use external sources for key components of the model, particularly substitution elasticities. In the agriculture literature, the estimation of input substitution elasticities has received considerable attention (e.g., Antle, Ball). However, these values have received relatively little attention in the forestry literature. Values as high as 2 to 3 have been used for the capital/labor (K/L) elasticity in timber production (Boyd and Daniels). These relatively large values were jus-

Table 5. Effect of Partial Equilibrium and Changing Substitution Elasticity Values in Land-Using Sectors on Output and Prices from TAX-2

Sector	Base Case	PE	Elasticities Used				
			Fixed Coeff.	Low-Level	Mid-Level	Cobb-Douglas	High-Level
Forestry K/L	1.5	1.5	0	0.3	0.5	1.0	1.5
K/D	1.0	1.0	0	0.3	0.5	1.0	1.5
AG-I K/L	0.7	0.7	0	0.3	0.5	1.0	1.5
K/D	0.3	0.3	0	0.3	0.5	1.0	1.5
AG-II K/L	0.7	0.7	0	0.3	0.5	1.0	1.5
K/D	0.4	0.4	0	0.3	0.5	1.0	1.5
AG-III K/L	0.7	0.7	0	0.3	0.5	1.0	1.5
K/D	0.3	0.3	0	0.3	0.5	1.0	1.5
Output							
Forestry	-1.42%	-3.10%	-2.60%	-1.60%	-1.48%	-1.19%	-1.10%
AG-I	-0.36%	-0.90%	-1.11%	-0.47%	-0.35%	-0.21%	-0.17%
AG-II	-0.22%	-0.30%	-0.54%	-0.27%	-0.22%	-0.16%	-0.14%
AG-III	0.35%	-0.15%	-0.25%	-0.26%	0.36%	0.47%	0.50%
Prices							
Forestry	1.73%	3.40%	3.21%	2.00%	1.76%	1.50%	1.40%
AG-I	0.20%	1.00%	1.32%	0.38%	0.19%	-0.01%	-0.08%
AG-II	-0.36%	0.60%	0.83%	-0.17%	-0.37%	-0.58%	-0.65%
AG-III	-0.89%	0.18%	0.16%	-0.72%	-0.90%	-1.08%	-1.14%
Land	-0.50%	—	4.80%	0.33%	-0.56%	-1.50%	-1.81%

tified by the long production period and ample substitution possibilities for timber growers.

In this study, a value of 1.5 was used for the *K/L* elasticity and 1.0 for the capital/land (*K/D*) elasticity in forestry. In agriculture a value of 0.7 was used for the *K/L* elasticity in all three sectors and 0.3 for the *K/D* elasticity in AG-I and AG-III and 0.4 in AG-II because of its larger scale. To test the sensitivity of the results to elasticity choice, the model was rerun by varying the elasticities in the land-using sectors from 0, the fixed coefficients case, 1.0, the Cobb-Douglas case, and an extreme value of 1.5.

Results from these experiments are shown in table 4. Because, these elasticity changes have only slight impacts on the rest of the economy, only the land-using sectors are shown. Resource inputs remain in place in the fixed coefficients case, so firms in the sector are unable to respond to changing input prices. As in the PE test, output declines and prices increase substantially in all of the land-using sectors because firms retain relatively low-valued excess capital. The extra capital in these sectors substantially increases the net price of land, thus increasing its marginal product value. As the elasticity of substitution increases, the adverse effects of TRA '86 caused by the decreasing net price of capital are ameliorated as firms buy relatively inexpensive labor. Although output still declines in three of

the four sectors, both the output declines and price increases are smaller than in the base case. Net land prices decline as firms liquidate low-valued capital causing a decline in the relatively fixed land factor.¹¹

This exercise emphasizes that the expected impacts from a policy change such as TRA '86 will be strongly determined by the choice of model parameters. The more easily that sectors can substitute inputs, the smaller are the direct and indirect impacts on that sector. In the same manner, if inputs are sticky in a sector, the impacts of policy changes may exceed predictions. For example, in much of the country, forestry is a residual land use. If forestry becomes submarginal for production, the land will revert to natural cover with less economic value rather than shifting to other sectors. While this land may provide other amenity values for the economy, it is no longer included in the productive base of the economy. Because our model does not allow for land to leave the economy, the full effect of TRA '86 on land resources in the economy may be understated.

¹¹ Sensitivity tests were also conducted with respect to export demand elasticities. Because prices in most land-based sectors rise, increasing the elasticity of export demand leads to a further decline in exports and production. Because international trade is only indirectly affected by the policy changes examined here, the changes in output are modest (less than 0.05% of land-based production).

Implications for Consumer Equity

The changes in net expenditures by income level after TRA '86 are shown in table 2. The restructuring of marginal tax brackets (*TAX-1*) has a positive direct effect on all income groups. High and low income classes are the biggest gainers because they receive the largest bracket reductions. TRA '86 also has indirect effects on consumer equity. Increases in capital taxation lower the price of both capital and land.¹² Landowners are hurt as a group, and taxpayers in lower income brackets are hurt the most (percentage-wise) because they own a substantial amount of land. By raising capital taxes differentially across sectors, the relative prices of consumer goods change.

Because each income class consumes commodities in differing proportions, we constructed a Laspeyres price index for each group, with the pre-tax quantities serving as quantity weights for all income groups. As shown in table 3, declining agricultural and housing costs give progressively lower relative benefits as the income class increases. Household "utility" changes, which quantify the combined effects of changes in purchasing power, available leisure, and the cost of living are also readily extracted from the MPS/GE computer output. These changes are also highly correlated with changes in income (table 3). Overall, the income and household "utility" effects are dominated by the direct effect due to changes in marginal bracket rates. When *TAX-2* is brought in, the income and utility gains from *TAX-1* to high income classes are severely dampened by the fall in capital prices and the resulting capital factor losses.

Conclusion

This study uses CGE techniques to assess the economy-wide changes expected from the TRA '86. The analysis focuses on the land-using sectors of the economy but considers changes in other sectors as well. Overall, the sectoral impacts are relatively small when compared to sector size, but the tax changes bring about important changes in land management. These land use impacts are often difficult to discern using

PE techniques because macro-policies such as TRA '86 change the relative prices of factor inputs and outputs causing more complicated production changes.

The forestry sector results of this study provide for interesting comparisons with previous PE analyses. Those studies (Guertin and Rideout, Greene and Kleunder) showed that forest management was adversely affected by TRA '86 and predicted large declines in timber production. Our analysis confirms these results in that the forest sector is relatively the hardest hit sector in the economy with output declining substantially. However, the decline is much less than has been shown previously because several general equilibrium considerations dampen TRA '86's effects. First, output prices increase because of increases in the relative cost of forest production and because of income-induced increases in the demand for sector outputs. Second, relatively cheaper labor substitutes for capital in the forest production process as firms move to more extensive land management practices. Finally, land is substituted for capital used in production as the value of capital gains decrease. In total, these market responses to the new economic situation cause a much smaller decline in the forest sector than has been projected.

The agriculture sector is also negatively affected by TRA '86. Because the loss of ITCs and other benefits is felt more strongly by the program crops sector, this part of the agriculture economy is the most strongly affected. The loss also affects the livestock sector through the rising price of feed grains. The vegetable and other crop category (AG-III), which is the least reliant on the ITCs, is the least affected by TRA '86, because it can absorb released low-cost capital and land to increase its production.

This general equilibrium analysis offers several general conclusions. First, unlike PE models of sectors such as forestry and agriculture, CGE analysis accounts for the finite resources and thereby avoids overestimating the effects of tax reform on sectoral output. Second, because CGE analysis provides a consistent accounting network, important changes in the demand for outputs are not overlooked. Third, when sectors within a CGE model have different inputs or receive different tax treatments (e.g., agriculture), valuable information will be lost if they are not disaggregated. Finally, a CGE model can explicitly explore equity issues by accounting for the costs and revenues occurring to different consumers.

¹² This result occurs because land is not free to move. Because the demand for agriculture goods declines and total agriculture revenue falls, the price of land falls relative to the price of labor.

[Received January 1989; final revision received May 1990.]

References

- Antle, J. M. "The Structure of U.S. Agricultural Technology, 1910-78." *Amer. J. Agr. Econ.* 66(1984):414-21.
- Ball, V. E. "Output, Input, and Productivity Measurement in U.S. Agriculture, 1948-79." *Amer. J. Agr. Econ.* 67(1985):475-86.
- Ballard, C. L., D. Fullerton, J. B. Shoven, and J. Whalley. *A General Equilibrium Model for Tax Policy Evaluation*. Chicago: University of Chicago Press, 1985.
- Boehlje, M., and H. Carman. "Tax Policy: Implications for Producers and the Agricultural Sector." *Amer. J. Agr. Econ.* 64(1982):1030-38.
- Boskin, M. J. "Taxation, Saving, and the Rate of Interest." *J. Polit. Econ.* 86(1978, part 2):S3-S27.
- Boyd, R. *An Economic Model of Direct and Indirect Effects of Tax Reform on Agriculture*. Washington DC: U.S. Department of Agriculture, Econ. Res. Serv. Tech. Bull. No. 1743, 1988.
- Boyd, R., and B. J. Daniels. "Capital Gains Treatment of Timber Income: Incidence and Welfare Implications." *Land Econ.* 61(1985):354-62.
- Caddy, V. "Empirical Estimation of the Elasticity of Substitution: A Review." Melbourne, Australia: Industries Assistance Commission Preliminary Work. Pap. OP-09, Impact Project, 1976.
- CCH Tax Law Editors. *State Tax Handbook*. Chicago: Commerce Clearing House, 1986.
- Fullerton, D., R. Gillett, and J. Mackie. "Investment Incentives Under the Tax Reform Act of 1986." *Compendium of Tax Research*, 1987. Washington DC: Internal Revenue Service, 1988.
- Greene, J. L., and R. A. Kleunder. "Effect of the 1986 Tax Reform Act on Forest Income and Values." *Proceedings of the 1988 Southern Forest Economics Workshop*, ed. R. C. Abt, 1988.
- Guertin, D. S., and D. B. Rideout. "The 1986 Tax Reform Act and Forest Investments: What Are the Consequences for Corporate Forestry?" *J. Forestry* 85(1987):29-31.
- Hansen, G. D., and D. R. Bertelsen. "Tax Reform Impacts on Agricultural Production and Investment Decisions." *Amer. J. Agr. Econ.* 69(1987):1013-20.
- Hertel, T. W., and M. E. Tsigas. "Tax Policy and U.S. Agriculture: A General Equilibrium Analysis." Dep. Agr. Econ. Work Pap. No. 87-2, Purdue University, Feb. 1987.
- Hertel, T. W., and M. E. Tsigas. "Tax Policy and U.S. Agriculture: A General Equilibrium Analysis." *Amer. J. Agr. Econ.* 70(1988):289-302.
- Houthakker, H. S. "An International Comparison of Household Expenditure Patterns, Commemorating the Century of Engle's Law." *Econometrica* 25(1957):532-51.
- Killingsworth, M. J. *Labor Supply*. New York: Cambridge University Press, 1983.
- Kokoski, M. F., and V. K. Smith. "A General Equilibrium Analysis of Partial-Equilibrium Welfare Measures: The Case of Climate Change." *Amer. Econ. Rev.* 77(1987):331-41.
- LeBlanc, M., and J. Hrubovcak. "The Effects of Tax Policy on Aggregate Agricultural Investment." *Amer. J. Agr. Econ.* 68(1986):767-77.
- Owen, J. D. "The Demand for Leisure." *J. Polit. Econ.* 79(1971):56-76.
- Rutherford, T. *General Equilibrium Modeling with MPS/GE*. Dep. Econ., University of Western Ontario, London, Ontario, March 1989.
- Stern, R. M., J. Francis, and B. Schumacher. *Price Elasticities in International Trade: An Annotated Bibliography*. London: Macmillan & Co., 1976.
- Summers, L. H. "Capital Taxation and Accumulation in a Life Cycle Growth Model." *Amer. Econ. Rev.* 71(1981):533-44.
- Tarr, D. G. "A General Equilibrium Analysis of the Welfare and Employment Effects of U.S. Quotas in Textiles, Autos, and Steel." Washington DC: Bureau of Econ. Staff Rep. to the Federal Trade Commission, 1984.
- U.S. Department of Agriculture, Economic Research Service. *Economic Indicators of the Farm Sector: Farm Sector Review, 1985*. ECIFS 5-4. Washington DC, 1985.
- . Various unpublished spreadsheets, 1986.
- U.S. Department of Commerce, Bureau of Economic Analysis. *Survey of Current Business*. Washington DC, May 1984, July 1985.
- . Various unpublished spreadsheets, 1986.
- U.S. Department of Energy, Energy Information Service. *Costs and Indices for Domestic Oil and Gas Field Equipment and Production Operations, 1985*. DOE/EIA-0185(85). Washington DC, 1986a.
- . *Petroleum Supply Annual, 1985*, 2 vol. DOE/EIA-0340(85). Washington DC, 1986b.
- U.S. Department of Labor, Bureau of Labor Statistics. *Consumer Expenditure Survey: Interview Survey, 1984*. Washington DC, 1986.

Appendix

Variable Descriptions for the Equations of the Model

Ψ	12 × 13 transformation matrix.
Ω_c	Defined in figure 1b.
τ_c	Marginal income tax rate on household c .
CD_c	Total government consumption by the c th household.
CD_j	Consumer demand of the i th product.
CE_j	Government endowment of product j .
DD_j	Demand for land in the j th industry.
DK_j	Demand for capital in the j th industry.
DL_j	Demand for labor in the j th industry.
EM_j	Demand elasticity of export demand.
FE_j	Endowment demand vector of adjusted elasticity of export demand.

GC_c	Gross consumption of household c .	SK_c	Supply of capital by the c th household, $c = 1, \dots, 6$.
GCE_i	Consumer demand of the i th consumer good, $i = 1, \dots, 13$.	SD_c	Supply of land by the c th household, $c = 1, \dots, 6$.
GD_j	Government demand for product j .	TC_c	Consumption taxes on the c th household.
GDD	Government demand for land.	TC_i	Excise tax on consumer good i .
GDK	Government demand for capital.	TD_j	Tax on land in the j th industry.
GDL	Government demand for labor.	TE	Total government endowments.
GE_j	Government endowment of product j .	TL_j	Tax on labor in the j th industry.
GSK_j	Government endowment of capital in the j th industry.	TK_j	Tax on capital in the j th industry.
INV_j	Investment in sector j .	TRN_c	Transfer payment to household c .
PIT_c	Personal income tax payment from household c .	TXO_j	Government output tax on the j th industry.
$\Sigma_L RAS_{jL}$	RAS-balanced I/O intermediate demands.	VM_j	Imports of product j .
RCS_{ic}	RAS-balanced matrix of each household's demand for each consumer good (6×13).	VX_j	Exports of product j .
SAV_c	Savings in household c .	Y_j	Total output of production sector $j = 1, \dots, 12$.
SL_c	Supply of labor by the c th household, $c = 1, \dots, 6$.	ZTA	Consumption + leisure coefficient (set to 1.5)

Consumer's Surplus Revisited

J. S. Shonkwiler

A technique for approximating welfare measures using knowledge of the Marshallian demand relation is developed and compared numerically to the true measures and other approximations. The technique does not rely on the solution of differential equations and is easily illustrated graphically. The technique should be useful as a pedagogic device and applicable in empirical work when derivation of the indirect utility function is burdensome.

Key words: compensating variation, consumer's surplus, Marshallian demand, welfare.

In 1976 R. D. Willig proposed a method to approximate compensating and equivalent variations when only the Marshallian demand relation was known. In fact, three distinct (though related) measures were introduced, each of which required information on the Marshallian consumer's surplus, the income elasticity, and the initial income level. One of these measures provided exact calculation of compensating variation, but only when the income elasticity was constant over the price change.

J. A. Hausman, five years later, correctly pointed out that approximation is unnecessary because a differential equation technique can be used to recover a local indirect utility function. Unfortunately, solution of the differential equation occasioned by a complicated Marshallian demand equation may be tedious in light of the relatively simple approximation proposed by Willig. Of course, it is tempting to suggest that empirical analyses should employ theoretically plausible demand specifications whose parameters can be related to a well-defined indirect utility function. Then neither approach need be adopted. This line of reasoning, however, ignores the occasional need for analytical expediency and the possibility that theoretically based systems may require modifications to better address complexities (see Pollack and Wales).

This reexamination provides an alternative technique for approximating welfare changes (compensating variation in particular) using a Marshallian demand relation. The technique is

about as easy to apply as Willig's rule-of-thumb measure and appears to be applicable over a comparable range of values of the income elasticity and consumer's surplus relative to income. Further, the approximation method has a distinct pedagogic attraction in terms of relating changes in income to changes in utility. The development proceeds by first discussing Willig's approximation and then developing the alternative technique. Some numerical comparisons then are provided.

The Two Techniques

Willig's rule of thumb measure for a single price change is

$$(1) \quad CV \approx A + \frac{\eta A^2}{2m^0},$$

where CV and A are compensating variation and consumer's surplus, η is the income elasticity and m^0 is the initial income level. The rule of thumb is derived from the successive solution of two differential equations and the application of a Taylor-series approximation. The derivation is explicitly developed in Willig's paper and will not be repeated here. Unfortunately, his approach is rather sterile and is not easy to explain graphically.

Some authors (e.g., Just, Hueth, and Schmitz) have tried to motivate the measure in (1) with an argument along the following lines:

Consider that $CV \approx A + 1/2 \Delta q \Delta p$, where p represents the price and q represents the quantity of a good. In figure 1, A is the area bounded by $p^0 abp'$ and CV is the area bounded by $p^0 acp'$. Δq is the line segment bc and $\Delta p = p' - p^0$. It

J. S. Shonkwiler is a professor of food and resource economics, University of Florida.

Florida Agricultural Experiment Station Journal Series No. R-00619.

Valuable comments by R. D. Emerson, J. W. Milon, and two anonymous reviewers are acknowledged by the author.

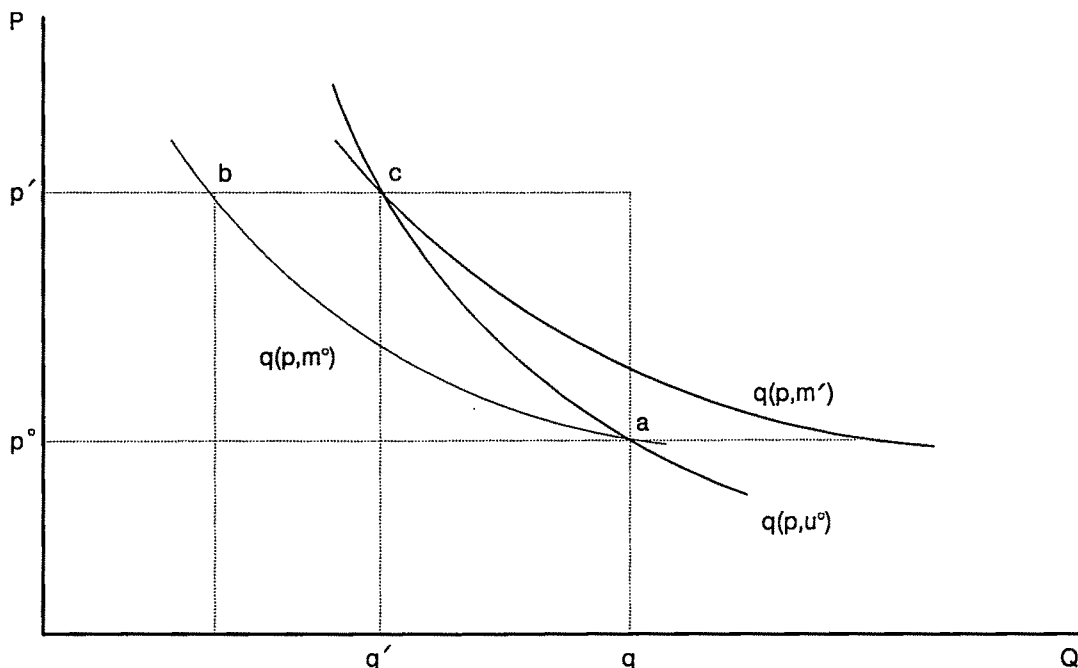


Figure 1. Compensated and uncompensated demand functions

is apparent that $\Delta q = q(p', m') - q(p', m^0)$. Then,

$$\lim_{m' \rightarrow m^0} \Delta q = \lim_{m' \rightarrow m^0} \frac{[q(p', m') - q(p', m^0)]}{m' - m^0} \cdot (m' - m^0),$$

or $dq = \lim_{\Delta m \rightarrow 0} \partial q / \partial m \Delta m$, so that

$$CV \approx A + 1/2 \frac{\partial q}{\partial m} \Delta m \Delta p.$$

Now, recognizing that $\eta = \partial q / \partial m \cdot m / q$, we have

$$(2) \quad CV \approx A + 1/2 \eta \frac{\Delta m}{m} q \Delta p.$$

At this point the argument is that for small Δp , both Δm and $q \Delta p$ may be approximated by A , hence the expression in (1). The problem with the final conclusion is that it is not entirely clear what Δm is measuring. If Δm provides a limit for CV , why not set Δm equal to CV and solve (2) for CV ? Alternatively, A underapproximates both CV and $q \Delta p$, suggesting (1) would not be nearly as good an approximation as it really is. Perhaps the unsatisfactory nature of the explanation for expression (1) should not be surprising because it is an improper simplification of

an approximation which originates from a much different approach.

The alternative technique for approximating CV has the advantages of not being based on differential equations and is easily developed graphically. In figure 1 note that the compensated demand curve approximately bisects the area between the two Marshallian demand curves when considering the price change from p^0 to p' . Recall that for a compensated demand function, $q(p, u^0)$, and an expenditure function, $C(p, u^0)$, having as arguments the reference level of utility, u^0 ,

$$(3) \quad \begin{aligned} CV &= \int_{p^0}^{p'} q(p, u^0) dp = \int_{p^0}^{p'} \frac{\partial C(p, u^0)}{\partial p} dp \\ &= C(p', u^0) - C(p^0, u^0) \\ &= m' - m^0 = \Delta m, \end{aligned}$$

and define two measures of consumer's surplus by

$$(4) \quad \begin{aligned} A &= \int_{p^0}^{p'} q(p, m^0) dp \text{ and} \\ A' &= \int_{p^0}^{p'} q(p, m') dp. \end{aligned}$$

This leads to the approximation

$$(5) \quad CV \approx A + 1/2(A' - A) = A + 1/2 \Delta A,$$

and using the definitions in (3) and (4) yields the expression

$$(6) \quad \Delta A = \int_{p^0}^{p'} [q(p, m^0 + \Delta m) - q(p, m^0)] dp.$$

Expression (6) illustrates that if the limits of integration are fixed, changes in income shift the Marshallian demand curve and thus lead to changes in consumer's surplus. Because m is a constant in the integration, standard results on differentiating under the integral sign (e.g., Woods, pp 141–42.) can be used such that

$$(7) \quad \lim_{\Delta m \rightarrow 0} \frac{\Delta A}{\Delta m} = \int_{p^0}^{p'} \frac{\partial q(p, m^0)}{\partial m} dp = \frac{\partial A}{\partial m}.$$

In terms of differentials this gives

$$(8) \quad CV \approx A + 1/2 \frac{\partial A}{\partial m} dm.$$

However, evaluation of (8) using information based solely on the Marshallian demand relation cannot be performed because dm , or its discrete approximation, Δm , is unknown. Yet, Δm is precisely the income change necessary to remain on the original compensated demand curve after the price change. Therefore, $\Delta m = CV$ and (8) can be solved in terms of CV to yield

$$(9) \quad CV \approx \frac{A}{1 - 1/2 \partial A / \partial m} = dS.$$

This approximation (dS) of compensated variation [approximate equivalent variation can be obtained by reversing the sign on the second term in the denominator on the right-hand side of (9)] follows directly from geometric arguments and does not require the solution of differential equations. As such, it is a more intuitive device for relating welfare measures to consumer's surplus. Interestingly, the dS measure has the same form in the linear specification as a measure suggested for expression (2) when Δm is equated to CV so that expression (2) becomes

$$(10) \quad \frac{A}{1 - 1/2 \frac{\partial q}{\partial m} (p' - p^0)}.$$

dS as an Approximation

The error bounds on the true value of CV for the dS approach can be shown to coincide ex-

actly with those presented by Willig. This is accomplished by writing expression (5) as an equality which depends on an unknown proportion ρ ,

$$(11) \quad CV = A + \rho \Delta A.$$

Some algebra leads to the result that

$$(12) \quad \rho = \frac{\Delta m - A}{\Delta A}.$$

Both Δm and ΔA depend on the unknown level of compensating income, m' . Willig provides bounds for m' , and these could be used to provide bounds for ρ , if desired. The discrete change form for dS becomes

$$(13) \quad CV = \frac{A}{1 - \rho \frac{\Delta A}{\Delta m}},$$

which identically equals Δm when expression (12) is substituted into (13). As would be expected, the bounds on CV are identical between the two approaches because $\Delta m = m' - m^0$ and the bounds on m' are used to provide the bounds on CV .

Two sources of error must be recognized in evaluating the approximation in (9): (a) the use of $1/2$ instead of the unknown proportion, denoted by ρ , and (b) the use of the derivative $\partial A / \partial m$ in place of the difference quotient $\Delta A / \Delta m$. As a consequence of these approximations, it appears that the dS measure does not perform well when the magnitude of compensating variation is large relative to income and the price elasticity and income elasticity are inelastic and elastic, respectively.

The results for a two good linear demand specification illustrate the effects of these factors on the approximation error. In table 1, various price, ϵ , and income elasticities of demand and corresponding levels of compensating variation are evaluated for a 100% price change at an initial income level of $m = 100$. Both W [Willig's rule-of-thumb measure from expression (1)] and dS are calculated from the initial price-quantity equilibrium. When the budget share, s , is large and when demand is price inelastic and income elastic, dS substantially overestimates CV . However, as the budget share and hence CV drops, the magnitude of the errors also declines.

These relationships are more systematically developed in table 2 for a log-linear demand function. Again, a 100% price change and an

Table 1. Linear Demand Function

ϵ (1) ^a	η (2)	s (3)	CV (4)	W (5)	dS (6)	% error of dS (7)
-.73	1.82	.88	158.59	84.51	280.00	76.55
-.4	1.5	.80	129.78	94.72	160.00	23.28
-.667	2.5	.48	67.91	44.80	80.00	17.80
-1.33	2.5	.48	43.02	19.20	40.00	7.02
-.80	4.8	.20	22.26	15.46	23.08	3.66
-.667	2.0	.24	21.35	18.56	21.05	1.41
-.545	.455	.44	35.87	34.33	35.56	.87

^a Where (1) is the price elasticity of demand, (2) is the income elasticity of demand, (3) is the budget share, (4) is actual compensating variation, (5) is Willig's rule-of-thumb measure, (6) is the alternative approximation of CV, and (7) is the percentage error of (6) relative to (4). Note the quantities in columns (1), (2), (3), (5) and (6) are calculated at the initial price and income levels. Initial income level is $m = 100$.

initial income level $m = 100$ are used. The results suggest that dS provides a reasonably accurate measure of CV provided calculated dS is not close to m , especially under price inelastic and income elastic demands.

The approximation value of dS is also illustrated for the quasi-homothetic demand specification. The budget share model that results from Deaton and Muellbauer's quasi-homothetic cost function is used as an example because A becomes a nonlinear function of m when the demand relation is written as a function of q . True CV can be calculated directly from the indirect utility function given in their article. The bounds on CV provided by Willig (W bounds) are also calculated. The results in table 3 again document that the dS measure provides reasonably good approximations to compensated variation only using information based on the Marshallian demand relation.

As a final consideration, deadweight loss (DWL) is measured by recognizing that q' in figure 1 can be estimated rather accurately in the approach developed. This occurs when the cal-

culated measure $m^\circ + dS$ is substituted for m° in the Marshallian demand relation. Then the approximate DWL becomes

$$(14) \quad DWL \approx dS - q(p', m^\circ + dS)\Delta p.$$

An example of the approximation's value in the context of the gasoline demand model presented by Hausman is considered. The calculated dS value is \$37.14 (versus a true CV of \$37.17) and approximate DWL is \$2.86 (versus a true value of \$2.88). This approach reinforces the view that knowledge of Marshallian demands can lead to good estimates of welfare measures.

Concluding Remarks

As long as the magnitude of the calculated measure dS is small relative to the initial income level the approximation performs on par with or better than Willig's rule-of-thumb measure. The degree of accuracy is also influenced by the relative magnitudes of the price and income elasticities of demand. Finally, it should be noted

Table 2. Log-Linear Demand Function

ϵ (1) ^a	η (2)	s (3)	CV (4)	W (5)	dS (6)	% error of dS (7)
-.35	1.85	.987	376.62	155.58	431.48	14.57
-.5	1.85	.802	166.02	107.25	172.32	3.79
-1.5	1.85	.200	13.17	13.02	13.17	.01
-.35	.85	.987	125.43	118.21	136.66	8.95
-.5	.85	.802	88.37	85.19	92.57	4.74
-1.5	.85	.200	12.35	12.33	12.36	.12
-.35	.35	.987	98.58	99.53	101.86	3.33
-.5	.35	.802	73.71	74.15	75.17	1.98
-1.5	.35	.200	11.98	11.98	11.99	.12

^aColumn definitions are identical to those in table 1.

Table 3. Results for Quasi-homothetic Demand Relation

ϵ (1) ^a	η (2)	s (3)	CV (4)	W (5)	dS (6)	W bounds (7)
-.84	.2	.75	32.072	32.283	32.435	32.228 to 32.435
-.60	1	.5	24.505	24.319	24.615	24.505 ^b
-.59	2	.25	12.203	12.154	12.246	12.195 to 12.302

^a Measures in columns (1) through (6) are identical to those in table 1.

^b Income elasticity is constant and bounds collapse to an exact measure.

that all results are applicable to the single price change case at the level of the individual consumer.

[Received January 1990; final version received May 1990.]

References

- Deaton, A., and J. Muellbauer. "An Almost Ideal Demand System." *Amer. Econ. Rev.* 70(1980):312-26.
- Hausman, J. A. "Exact Consumer's Surplus and Dead-weight Loss." *Amer. Econ. Rev.* 71(1981):662-76.
- Just, R. E., D. L. Hueth, and A. Schmitz. *Applied Welfare Economics and Public Policy*. Englewood Cliffs NJ: Prentice-Hall, 1982.
- Pollack, R. A., and T. J. Wales. "Demographic Variables in Demand Analysis." *Econometrica* 49(1981):1533-51.
- Willig, R. D. "Consumer's Surplus Without Apology." *Amer. Econ. Rev.* 66(1976):589-97.
- Woods, F. S. *Advanced Calculus*. Boston: Ginn and Co. 1934.

The Calculation of Research Benefits with Linear and Nonlinear Specifications of Demand and Supply Functions

Jan P. Voon and Geoff W. Edwards

This paper provides a comparison of research benefits for linear and nonlinear constant elasticity (NLCE) specifications of supply and demand with a pivotal supply shift framework. The comparison allows for a sensitivity test of results to different values for demand and supply elasticity. The main finding is that the values of the gross annual research benefits calculated for the NLCE specification are larger than those calculated for the linear specification with price-elastic supply but are smaller with price-inelastic supply. The analysis suggests that the use of an NLCE specification and a pivotal shift due to research is usually preferable to use of a linear supply curve with pivotal shift.

Key words: demand-supply specifications, elasticities, model comparison, research benefits.

Market models have been used extensively for measuring the level and the distribution of benefits from R&D-induced supply shifts. In those models the supply and demand schedules were assumed to be either linear or nonlinear with constant elasticity (NLCE), and the R&D-induced supply shifts were assumed to be either parallel or nonparallel. Comparisons of research benefits occurring with linear parallel and linear nonparallel shifts in supply were provided by Lindner and Jarrett, Rose, Wise and Fell, and Norton and Davis. Comparisons of research benefits for linear and nonlinear specifications of the demand and supply curves have not been provided. The main aim in this paper is to provide such comparisons.

Both the linear and the NLCE frameworks have frequently been used in empirical work on evaluation of research benefits. The use of either framework involves little or no computational difficulty and appears practical for most empirical analyses. The NLCE pivotal supply shift framework has been adopted by Peterson; Ayer and Schuh; Akino and Hayami; Flores-Moya, Evenson, and Hayami; Nagy and Furtan; Wise; and Zentner and Peterson, among others. The

linear pivotal supply shift framework can be found in Lindner and Jarrett; Rose; McLean; Wise; and Norton, Ganoza, and Pomareda. Comparison of research benefits for linear and NLCE frameworks provides an understanding of the differences in results caused by model specification. This understanding is important in choosing between frameworks for estimating economic benefits from research and in interpreting results. The key point is that if a particular model specification is a better description of reality, then use of the alternative model specification can cause overestimation (or underestimation) of returns to research.

Comparison of the measured economic benefits from research using linear and NLCE supply curves requires specification of the type of supply shift resulting from research. This is not a straightforward task. No NLCE counterpart exists for some specifications of shifts in linear supply curves. The best example of this is the parallel shift in a linear supply curve, a specification used by several researchers (e.g., Rose, Edwards and Freebairn). Researchers using NLCE supply curves, on the other hand, have not used parallel shifts in supply due to research. This may be the result of an absence of analytical techniques for specifying a parallel shift of the NLCE supply curve and the function of

Jan P. Voon and Geoff W. Edwards are, respectively, a graduate student and a senior lecturer at LaTrobe University, Australia.

the new supply curve because NLCE supply curves always pass through the origin.

Because our aim is to compare estimated research benefits with linear and NLCE supply curves, it is important to formulate the analysis so that other factors affecting the size of research benefits, including the nature of the supply shift, are identical for the linear and NLCE cases. In order to do this, our approach confines the comparison to estimates of research benefits resulting from corresponding pivotal shifts in linear and NLCE supply curves. Because NLCE supply curves always pass through the origin, the pivotal shift for these curves is always about the origin.¹ For the linear case, however, the intercept is a function of the elasticity of supply. For supply elasticity greater than unity, the linear supply curves pivot on the positive price axis. For supply elasticity equal to unity, the linear supply curves pivot on the origin. For supply elasticity smaller than unity, the linear supply curves pivot on the negative price axis. The effects of the use of different values for the elasticity of supply on the size and distribution of research benefits are examined in this paper.

Research Benefits for Linear and Nonlinear Pivotal Supply Shifts

In this section, we outline the procedures used for comparing the level and the distribution of research benefits using linear and nonlinear constant elasticity specifications of the demand and supply curves. For comparison, we fix the following conditions [fig. 1(a) and 1(b)]: the supply shifts are pivotal for both cases; shifts in supply due to research, measured vertically as cost reductions at the initial equilibria, are identical ($A - D$) for the two cases; price (P_0) is identical for the two cases at the initial equilibrium, as is quantity (Q_0); and price elasticities of demand are identical for the two cases at the initial equilibrium, as are price elasticities of supply.

The general functional form for linear and NLCE demand and supply curves may be represented by $Q = f(P)$. When solving $Q = f(P)$ for P and Q , the notation $P = f^{-1}(Q)$ is used. That is, f^{-1} is the function inverse to f . The linear inverse demand curve, in an algebraic form,

¹ The power function of the nonlinear supply curve specified by Miller, Rosenblatt, and Hushak can pass through other points on the price axis, but this particular functional form is computationally difficult and is not practical for empirical purposes (e.g., a numerical calculation of producer surplus).

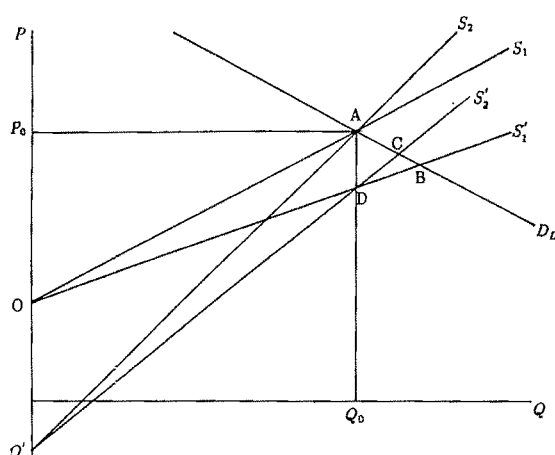


Figure 1(a). Effect of research with linear demand and supply curves

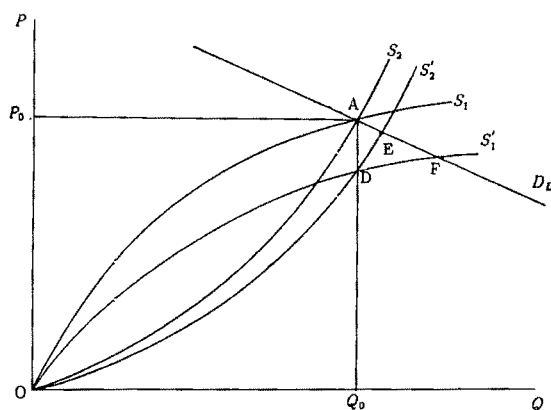


Figure 1(b). Effect of research with nonlinear constant elasticity demand and supply curves

is represented by $P = a - \alpha Q$, where a is a positive constant and α is the slope of the demand curve. Note that $\alpha = P/\eta Q$, where η denotes the own price elasticity of demand at the initial equilibrium. The linear inverse supply curve in the absence of research is represented by $P = b + \beta Q$, where b is a constant term and β is the slope of the supply curve. Again, $\beta = P/eQ$, where e is the own-price elasticity of supply at the initial equilibrium. The linear supply curve with research is denoted by $P' = b + \beta' Q'$, where β' is the slope of the 'with research' supply curve. Note that $\beta' = P(1/e - k)/Q$, where k [$k = AD/P_0$, as illustrated in figure 1(a) and 1(b)] represents the proportionate vertical shift in the supply curve ($0 < k < 1$).

The constant elasticity demand curve is represented by $P = AQ^{-\sigma}$, where A is a positive constant and σ represents the price flexibility of demand ($\sigma = 1/\eta$). The constant elasticity supply curve without research is $P = BQ^\gamma$, where B is a positive constant and γ is the price flexibility of supply ($\gamma = 1/e$) and the constant elasticity supply curve with research is $P' = (1 - k)BQ^\gamma$. The equilibrium price and quantity without the supply shift are P_0 and Q_0 .

Comparison of the gross annual research benefits (GARB) is performed for area OAB (or area $O'AC$) in figure 1(a) and area OAE (or area OAF) in figure 1(b).² The distribution of research benefits between producers and consumers with the linear and NLCE specifications is also compared.

Four combinations of the linear and NLCE demand and supply curves are identified. These are case 1: linear demand and supply curves ($D_L S_L$), case 2: NLCE demand and linear supply curves ($D_N S_L$), case 3: linear demand and NLCE supply curves ($D_L S_N$), and case 4: NLCE demand and supply curves ($D_N S_N$).

Let TS ($TS = GARB$) be the change in total economic surplus, CS be the change in consumer surplus, PS be the change in producer surplus, and subscripts 1, 2, 3, and 4 denote cases 1, 2, 3, and 4, respectively.³ The formula for calculating the level and the distribution of research benefits for case 1 [fig. 1(a)] is expressed as follows:

$$(1) \quad TS_1 = 1/2kP_0Q_1,$$

$$(2) \quad CS_1 = 1/2(P_0 - P_1)(Q_0 + Q_1),$$

² Consider first the linear case as shown in figure 1(a). For $e < 1$, the gross annual research benefit equals area $O'AC$ and for $e > 1$, the gross annual research benefit equals area OAB . Note that area OAB is bigger than area $O'AC$ by area CBD . For identical cost reductions, the values of the gross annual research benefits are little affected by changes in supply elasticity (since area CBD is very small) notwithstanding that, for $e < 1$, the linear supply curve passes through the fourth quadrant negative euclidean space. Now consider the NLCE case as shown in figure 1(b). For $e < 1$, the gross annual research benefit is represented by area OAE and for $e > 1$, the gross annual research benefit is represented by area OAF . NLCE supply curve takes a convex shape for $e < 1$ and a concave shape for $e > 1$.

³ The advantages and limitations of consumer and producer surplus measures are common to the linear and NLCE specifications that we are considering. Consumer surplus measures are widely used for analyzing welfare effects of price changes for agricultural products. This is generally regarded as appropriate, largely because income effects caused by price changes are likely to be small since consumers spend a very small fraction of their income on a particular food item (Bigman and Shalit). Producer surplus, also used widely in welfare analysis, is open to more serious questioning. Use of producer surplus is most clearly appropriate when rents accrue to a single-fixed factor, all other factors being in perfectly elastic supply (Mishan).

$$(3) \quad PS_1 = TS_1 - CS_1.$$

The "with research" equilibrium point for case 1 can be determined using the alternative formulas developed in this paper:⁴

$$(4) \quad P_1 = P_0[1 - ke/(e + \eta - ke\eta)],$$

$$(5) \quad Q_1 = Q_0[1 + ke\eta/(e + \eta - ke\eta)].$$

The formula for calculating research benefits with NLCE demand and supply (case 4) is expressed as

$$(6) \quad TS_4 = \int_0^{Q_0} S(dQ) - \int_0^{Q_0} S'(dQ) + \int_{Q_0}^{Q_1} D(dQ) - \int_{Q_0}^{Q_1} S'(dQ),$$

$$(7) \quad CS_4 = \int_{P_1}^{P_0} D(dP),$$

$$(8) \quad PS_4 = TS_4 - CS_4,$$

where $Q_1 = \exp(-\ln(1 - k)/(\sigma + \gamma))$ and $P_1 = (Q_1)^{-\sigma}$.

Similarly, changes in research benefits for the intermediate cases (i.e., cases 2 and 3) can be calculated using the above approaches.

Results

For illustrative purposes, the initial equilibrium prices and quantities were set at unity for each of the four cases, k was set at 0.1 (so the absolute shift is 0.1 P_0), and a range of demand and supply elasticities was used for comparison.

The size of $GARB$ and the calculated distribution of research benefits for case 2 ($D_N S_L$) are

⁴ Past models (Norton, Ganoza, and Pomareda; Rose) used the equations of Pinstrup-Andersen, Ruiz de Londoño, and Hoover (where $P_1 = P_0[1 - ke/(e + \eta)]$ and $Q_1 = Q_0[1 + ke\eta/(e + \eta)]$) for deriving the post-innovation (or "with research") equilibrium point [i.e., point B or C in fig. 1(a)], but these equations, as pointed out by Rose, are strictly correct only if the supply shift is parallel. For the linear pivotal supply shift model, the actual equations for determining the "with research" equilibrium point have been developed in this paper [i.e., equations (4) and (5)]. When research benefits calculated using the past approach were recalculated using the actual approach (using the parameter values assumed in tables 1 and 2 of this note), the values of $GARB$ were observed to be only marginally higher. The difference was found to be approximately 0%–3% with demand and supply elasticities in the range 0.1 and 5.0. A larger difference, however, is found in the calculated distributions of research benefits. For instance, recalculations showed that consumers' share of research benefits calculated from the past approach is smaller than those calculated from the actual approach by 10%–30% with larger values of η and e . Further details are available from the authors.

similar to those for case 1 ($D_L S_L$) for each combination of η and e , while the size of $GARB$ and the calculated distribution of research benefits for case 3 ($D_L S_N$) correspond closely to those for case 4 ($D_N S_N$). That is, for the marginal changes under consideration, the choice between the linear and nonlinear specifications of demand does not have a significant influence on price of the commodity "with research." It, therefore, has little influence on the size or distribution of $GARB$. Hence, for comparison, the results for cases 1 and 3 are taken (see table 1). The results for cases 2 and 4 are excluded from table 1 because the size and distribution of research benefits for case 2 is similar to that for case 1, while the size and distribution of research benefits for case 4 approximate that for case 3.

For supply elasticities less than about unity, the values of $GARB$ calculated for case 1 ($D_L S_L$) are higher than those calculated for case 3 ($D_L S_N$). For elasticities of supply greater than about unity, however, the values of $GARB$ are larger for case 3 than case 1. The magnitudes of the differences in $GARB$ for these two cases are larger the smaller the values for the supply elasticities (e.g., a difference of about 142% for $e = 0.25$ and about

16% for $e = 0.75$)⁵ (see table 2). The differences in $GARB$ are very small when the price elasticity of supply approaches unity.

For each of the four cases, the values of $GARB$ are insensitive to the value of the price elasticity of demand. The values of $GARB$ calculated using the $D_L S_L$ and $D_N S_L$ specifications (cases 1 and 2) are not sensitive to the choice of supply elasticity. However, the values of $GARB$ calculated using the $D_L S_N$ and $D_N S_N$ specifications (cases 3 and 4) are extremely sensitive to changes in supply elasticity. The explanation for this is as follows: for $e < 1$, the NLCE supply curves slope upwards at an increasing rate [see S_2 and S'_2 in fig. 1(b)], whereas for $e > 1$, the NLCE supply curves slope upwards but at a decreasing rate [see S_1 and S'_1 in fig. 1(b)]. Given an identical size of cost reduction at the initial equilibrium, the "with research" and the "without research" supply curves for $e < 1$ lie closer together (i.e., the mean vertical distance between the two supply curves S_2 and S'_2 is smaller), resulting in smaller values of $GARB$ relative to the linear

⁵ This is for $|\eta| = 1$. The pattern is similar for other values of $|\eta|$.

Table 1. Comparison of the Level and the Distribution of Research Benefits for the $D_L S_L$ (Case 1) and $D_L S_N$ (Case 3) Specifications

η	e	Case 1 ($D_L S_L$)			(\$)	Case 3 ($D_L S_N$)		
		CS	PS	TS		CS	PS	TS
-0.01	0.01	0.0500	0.0000	0.0500	0.0513	-0.0503	0.0010	
	0.25	0.0963	-0.0462	0.0501	0.0964	-0.0763	0.0201	
	0.75	0.0988	-0.0488	0.0501	0.0988	-0.0559	0.0429	
	1.00	0.0992	-0.0491	0.0501	0.0991	-0.0491	0.0501	
	2.00	0.0997	-0.0496	0.0501	0.0996	-0.0329	0.0667	
-0.25	0.01	0.0039	0.0462	0.0501	0.0041	-0.0030	0.0010	
	0.25	0.0510	0.0003	0.0506	0.0517	-0.0310	0.0207	
	0.75	0.0772	-0.0262	0.0510	0.0767	-0.0329	0.0438	
	1.00	0.0825	-0.0315	0.0510	0.0817	-0.0307	0.0510	
	2.00	0.0919	-0.0408	0.0511	0.0905	-0.0227	0.0678	
-1.00	0.01	0.0010	0.0491	0.0501	0.0010	0.0000	0.0010	
	0.25	0.0206	0.0304	0.0510	0.0211	0.0000	0.0211	
	0.75	0.0458	0.0065	0.0522	0.0452	-0.0001	0.0451	
	1.00	0.0540	-0.0014	0.0526	0.0527	-0.0001	0.0526	
	2.00	0.0740	-0.0204	0.0536	0.0703	-0.0002	0.0701	
-20.00	0.01	0.0001	0.0500	0.0501	0.0001	0.0010	0.0010	
	0.25	0.0013	0.0500	0.0513	0.0013	0.0200	0.0213	
	0.75	0.0041	0.0499	0.0539	0.0040	0.0429	0.0468	
	1.00	0.0055	0.0497	0.0553	0.0053	0.0500	0.0553	
	2.00	0.0124	0.0488	0.0611	0.0105	0.0664	0.0770	

Note: Negative (-) indicates a loss in surplus.

Table 2. Differences in Values of *GARB* Calculated Using the $D_L S_L$ (Case 1) and the $D_L S_N$ (Case 3) Specifications

Supply Elasticity	Case 1 ($D_L S_L$)	Case 3 ($D_L S_N$)	Difference in <i>GARB</i> ^a
	----- (\$)	-----	(%)
0.01	0.05005	0.00104	4700
0.25	0.05102	0.02107	142
0.50	0.05172	0.03512	47
0.75	0.05224	0.04515	16
1.00	0.05264	0.05268	0
1.25	0.05294	0.05853	-11
1.50	0.05319	0.06322	-19
2.00	0.05357	0.07024	-31
5.00	0.05455	0.08780	-61

^a Major finding from table 1, using $\eta = -1$.

supply curves. The "with research" and the "without research" supply curves for $e > 1$ lie further from each other (i.e., the mean vertical distance between the two supply curves S_1 and S'_1 is greater), resulting in larger values of *GARB* relative to the linear supply curves.

Consumers' benefits from research are similar for each of the four cases. The significant difference in *GARB* between cases 1 and 2 on the one hand and cases 3 and 4 on the other reflect differences in benefits to producers. Producers can lose for each of the four cases when the demand for a commodity is price inelastic (i.e., $|\eta| < 1$), and with nonlinear supply (cases 3 and 4) producers gain from research only when $|\eta| > 1$.⁶ This contrasts with a linear parallel shift in supply when producers always gain (unless demand is perfectly inelastic). For each of the four cases, producers' gains from research increase (or losses decrease) with increases in η and fall with increases in e .

Concluding Comments

The major finding reported in this note is that the values of *GARB* calculated using the linear pivotal supply shift model (cases 1 and 2) are substantially larger than those calculated using the constant elasticity pivotal supply shift model

(cases 3 and 4) when the price elasticity of supply for the commodity is significantly lower than unity. However, when the elasticity of commodity supply is greater than one, the values of *GARB* calculated using the linear supply shift framework are considerably smaller than those calculated using the constant elasticity supply shift framework. An important implication of this finding is as follows: if the constant elasticity specification is a better description of reality, the linear pivotal supply shift model for evaluating research benefits leads to marked overestimation of research benefits when the commodity supply is inelastic but to considerable underestimation when supply is elastic.

It is debatable whether the real-world situation corresponds more closely to constant elasticity or to constant slope. In our view, commodity supply curves in agriculture are more likely to take the constant elasticity form. Both short-run and the long-run supply for many rural commodities have often been reported in the inelastic range (less than unity) (e.g., Vincent, Powell, and Dixon; Tweeten). With inelastic supply at the initial equilibrium, a linear supply curve passes through the negative fourth quadrant. An extrapolation of the linear inelastic commodity supply curve to the negative fourth quadrant (i.e., production of the commodity at undefined negative prices) is unrealistic. This problem is avoided by the use of constant elasticity supply curves; these curves pass through the origin. Our analysis, therefore, leads us to the view that the use of a nonlinear, constant elasticity specification of the supply curve and a pivotal shift due to research is usually preferable to use of a linear supply curve with pivotal shift.

[Received January 1990; final revision received August 1990.]

References

- Akino, M., and Y. Hayami. "Efficiency and Equity in Public Research: Rice Breeding in Japan's Economic Development." *Amer. J. Agr. Econ.* 57(1975):1-10.
- Ayer, H. W., and G. E. Schuh. "Social Rates of Return and Other Aspects of Agricultural Research: The Case of Cotton Research in São Paulo, Brazil." *Amer. J. Agr. Econ.* 54(1972):557-69.
- Bigman, D., and H. Shalit. "Applied Welfare Analysis for Consumers with Commodity Income." *De Economist* 131(1983):31-45.
- Edwards, G. W., and J. W. Freebairn. "The Gains from Research into Tradable Commodities." *Amer. J. Agr. Econ.* 66(1984):41-49.
- Flores-Moya, P., R. E. Evenson, and Y. Hayami. "Social

⁶ Our results are consistent with the major finding reported in Miller, Rosenblatt, and Hushak. That paper focused on the effects of supply shifts on producers' surplus. For supply and demand curves that are linear or power functions, a small downward pivot of the supply curve increased producers' surplus only if the equilibrium point was far enough into the elastic region of the demand curve, and any downward pivot of the supply curve decreased producers' surplus if the equilibrium point was in the inelastic region of the demand curve.

- Returns to Rice Research in the Philippines: Domestic Benefits and Foreign Spillover." *Econ. Develop. and Cultur. Change* 26(1978):591-607.
- Lindner, R. K., and F. G. Jarrett. "Supply Shifts and the Size of Research Benefits." *Amer. J. Agr. Econ.* 60(1978):48-56.
- McLean, I. W. "The Demand for Agricultural Research in Australia 1870-1914." *Australian Econ. Pap.* 21 (1982):294-308.
- Miller, G. Y., J. M. Rosenblatt, and L. J. Hushak. "The Effects of Supply Shifts on Producers' Surplus." *Amer. J. Agr. Econ.* 70(1988):886-91.
- Mishan, R. J. "What Is Producer's Surplus?" *Amer. Econ. Rev.* 58(1968):1269-82.
- Nagy, J. G., and W. H. Furtan. "Economic Costs and Returns from Crop Development Research: The Case of Rapeseed Breeding in Canada." *Can. J. Agr. Econ.* 26(1978):1-14.
- Norton, G. W., and J. S. Davis. "Evaluating Returns to Agricultural Research: A Review." *Amer. J. Agr. Econ.* 63(1981):686-99.
- Norton, G. W., V. G. Ganoza, and C. Pomareda. "Potential Benefits of Agricultural Research and Extension in Peru." *Amer. J. Agr. Econ.* 69(1987):247-57.
- Peterson, W. L. "Returns to Poultry Research in the United States." *J. Farm Econ.* 49(1967):656-69.
- Pinstrup-Andersen, P., N. R. De Londoño, and E. Hoover. "The Impact of Increasing Food Supply on Human Nutrition: Implication of Commodity Priorities in Agricultural Research and Policy." *Amer. J. Agr. Econ.* 58(1976):131-42.
- Rose, F. "Supply Shifts and the Size of Research Benefits: Comment." *Amer. J. Agr. Econ.* 62(1980):834-35.
- Tweeten, L. G. *Foundations of Farm Policy*. Lincoln: University of Nebraska Press, 1970.
- Vincent, D. P., A. A. Powell, and P. B. Dixon. "Changes in Supply of Agricultural Products." *Agriculture in the Australian Economy*, ed. D. B. Williams. Sydney: Sydney University Press, 1982.
- Wise, W. S. "The Shift of Cost Curves and Agricultural Research Benefits." *J. Agr. Econ.* 35(1984):21-30.
- . "The Theory of Agricultural Research Benefits." *J. Agr. Econ.* 32(1981):147-56.
- Wise, W. S., and E. Fell. "Supply Shifts and the Size of Research Benefits: Comment." *Amer. J. Agr. Econ.* 62(1980):838-40.
- Zentner, R. P., and W. L. Peterson. "An Economic Evaluation of Public Wheat Research and Extension Expenditures in Canada." *Can. J. Agr. Econ.* 32(1984):327-53.

Dairy Farm Efficiency Measurement Using Stochastic Frontiers and Neoclassical Duality

Boris E. Bravo-Ureta and Laszlo Rieger

This paper presents a stochastic efficiency decomposition model based on Kopp and Diewert's deterministic methodology. The stochastic model is used to analyze technical, economic, and allocative efficiency for a sample of New England dairy farms. The results suggest that mean economic efficiency for the farmers in the sample is about 70% and that, on average, there is little difference between technical (83.0%) and allocative (84.6%) efficiency. Analyses of the relationship between efficiency and four socioeconomic variables—farm size, education, extension, and experience—reveal that, despite some statistically significant associations, efficiency levels are not markedly affected by these variables.

Key words: economic, technical, and allocative efficiency; milk production; stochastic frontiers.

Efficiency measurement has received considerable attention from both theoretical and applied economists. From a theoretical point of view, there has been a spirited exchange about the relative importance of the various components of firm efficiency (Leibenstein 1966, 1978; Co-manor and Leibenstein; Stigler). From an applied perspective, measuring efficiency is important because this is the first step in a process that might lead to substantial resource savings. These resource savings have important implications for both policy formulation and firm management.

In the policy arena, there is a continuing controversy regarding the connection between farm size, efficiency, and the structure of production agriculture (e.g., Hall and LeVeen). For individual farms, gains in efficiency are particularly important in periods of financial stress similar to that experienced in U.S. agriculture in the 1980s. Efficient farms are more likely to generate higher incomes and thus stand a better chance of surviving and prospering.

The current interest in efficiency measurement finds its origin in a pioneering paper published by M. J. Farrell over thirty years ago. The approach proposed by Farrell distinguishes between technical and allocative efficiency where the former refers to the ability of producing a given level of output with a minimum quantity of inputs, given technology; the latter refers to the choice of the optimal input proportions given relative prices. Economic or total efficiency is the product of technical and allocative efficiency. Farrell's model, which is known as a deterministic nonparametric frontier (Forsund, Lovell, and Schmidt), attributes any deviation from the frontier to inefficiency and imposes no functional form on the data. Several extensions of Farrell's deterministic model have been made by Aigner and Chu, Timmer, Afriat, Richmond, Schmidt (1976), and Greene, among others.

A deficiency characterizing all deterministic frontier models is their sensitivity to extreme observations. A more recent approach for measuring efficiency, which seeks to ameliorate the extreme observation problem, is the stochastic frontier model developed by Aigner, Lovell, and Schmidt, and by Meeusen and van den Broeck. The stochastic frontier model assumes an error term with two additive components—a symmetric component which accounts for pure random factors, and a one-sided component which captures the effects of inefficiency relative to the

Boris E. Bravo-Ureta is an associate professor, and Laszlo Rieger is a graduate research assistant, both in the Department of Agricultural and Resource Economics, University of Connecticut.

Scientific Contribution No. 1274 of the Storrs Agricultural Experiment Station, University of Connecticut.

The authors are grateful to Emilio Pagoulatos, Linda Lee, Ricardo Quiroga, Stan Seaver, T. C. Lee, Richard Nehring, and *Journal* referees for their comments on earlier drafts of this paper.

stochastic frontier. A recent extension by Jon-drow et al. has solved the previous inability of deriving individual firm efficiency measures from stochastic frontiers.

Farrell's model allows the computation of allocative, technical and, hence, of economic efficiency, but this computation is restricted to a technology exhibiting constant returns to scale. Recent work by Kopp and Diewert, Akridge, Kumbhakar, and Schmidt and Lovell has led to alternative formulations of parametric models which relax the linear homogeneity restriction while enabling the calculation of the various efficiency indexes.¹ In spite of these methodological developments, efficiency studies dealing with agriculture, in general, and U.S. milk production, in particular, have focused primarily on technical efficiency.

This paper contributes to the literature on firm-level efficiency measurement by extending Kopp and Diewert's decomposition technique from a deterministic to a stochastic model. This stochastic formulation yields technical, economic, and allocative efficiency measures that are free from distortions, stemming from statistical noise, inherent in deterministic models. In addition, the model makes possible a comprehensive efficiency analysis relying only on the econometric estimation of a production frontier, which is helpful because the firm-level price data required to estimate dual (cost or profit) models are often unavailable or inadequate (Quiggin and Bui-Lan). As far as we know, this is the only study that analyzes technical, economic, and allocative inefficiency besides the work by Taylor, Drummond, and Gomes for a sample of Brazilian farmers. The latter paper, however, utilized a deterministic model and, consequently, the resulting efficiency measures are sensitive to outliers.

The stochastic efficiency decomposition model is applied to a sample of New England dairy farms in order to quantify technical, allocative, and economic efficiency levels. The results show that focusing only on technical efficiency, as the applied studies of dairy farm efficiency have done (e.g., Bravo-Ureta, Grisley and Mascarenhas, Quiroga and Elterich, and Tauer and Belbase) considerably understates the potential gains from improvements in overall performance. An eval-

uation is made also of the relationships between inefficiency and farmers' socioeconomic characteristics. Efficiency in dairy production is important because technological innovations (e.g., bovine somatotropin, or bST), the seemingly chronic oversupply of dairy products, and attempts to reduce fiscal outlays for agricultural programs suggest that the trend toward fewer milk producers is likely to continue. During this period of adjustment, efficiency will be a key determinant of dairy farm survival.

Analytical Framework

Following Kopp and Diewert, we begin by assuming that the production frontier is given by

$$(1) \quad Q = g(X_a),$$

where Q is output, and X_a is a vector of variable inputs. The technically efficient input vector (X_t) for a given level of output (\bar{Q}) is derived by solving simultaneously equation (1) and the input ratios $X_1/X_i = k_i (i > 1)$, where k_i is the ratio of observed inputs X_1 and X_i at output \bar{Q} .²

Assuming that the production frontier is self-dual (e.g., Cobb-Douglas), then the corresponding cost frontier, derived analytically, can be written in general form as

$$(2) \quad C = h(P, Q),$$

where C is the minimum cost associated with the production of output Q , and P is a vector of input prices.³ Applying Shephard's lemma, we obtain

$$(3) \quad \frac{\partial C}{\partial P_i} = X_i(P, Q),$$

which is the system of minimum cost input demand equations. Substituting a firm's input prices and output quantity into the demand system in equation (3), we obtain the economically efficient input vector (X_e). The X_t and X_e vectors can be used to compute the cost of the technically efficient ($X_t'P$) and the economically efficient ($X_e'P$) input combinations associated with the firm's observed output. In addition, the cost of the firm's actual operating input combination

² For a homothetic production function, this solution essentially corresponds to the point where the efficient isoquant for \bar{Q} intersects the isocline $X_1/X_i = k_i$.

³ The method developed by Kopp and Diewert can also be implemented by estimating a cost frontier which need not be self-dual (e.g., translog). However, given the data available to us, we follow the primal route and use a self-dual function.

¹ The models developed by Akridge, Kumbhakar, and by Schmidt and Lovell require the joint estimation of the production frontier and the first-order conditions for profit maximization or cost minimization. Therefore, the high data requirements of these models preclude their consideration in this paper.

is given by $X_a'P$. These three cost measures can now be used to compute technical (TE) and economic (EE) efficiency indexes as follows:

$$(4) \quad TE = (X_i'P)/(X_a'P),$$

$$(5) \quad EE = (X_e'P)/(X_a'P).$$

Finally, allocative efficiency (AE), derived from equations (4) and (5), is given by

$$(6) \quad AE = (EE)/(TE) = (X_e'P)/(X_i'P).$$

Kopp and Diewert's decomposition approach is based on a deterministic frontier which imposes the limiting assumption that any deviation from the frontier is the result of inefficiency; hence, the resulting inefficiency measures are biased (Schmidt, 1985-86). To avoid this problem, we estimate a stochastic production frontier model and use the approach introduced by Jondrow et al. to purge the purely random error from the efficiency component.

To illustrate how the random error is purged, consider the stochastic production frontier,

$$(7) \quad Q = f(X_a) + \epsilon,$$

where

$$(8) \quad \epsilon = v - u$$

is the composed error term (Aigner, Lovell, and Schmidt; Meeusen and van den Broeck). The two components v and u are assumed to be independent of each other, where v is the two-sided, normally distributed random error ($v \sim N(0, \sigma_v^2)$), and u is the one-sided efficiency component with a half-normal distribution ($u \sim |N(0, \sigma_u^2)|$). The maximum likelihood estimation of equation (7) provides estimators for β , λ , and σ^2 , where β was defined earlier, $\lambda = \sigma_u/\sigma_v$, and $\sigma^2 = \sigma_u^2 + \sigma_v^2$. Given the assumptions on the distribution of v and u , Jondrow et al. show that the conditional mean of u given ϵ is equal to

$$(9) \quad E(u | \epsilon) = \sigma_* \left(\frac{f^*(\lambda\epsilon/\sigma)}{1 - F^*(\lambda\epsilon/\sigma)} - \frac{\lambda\epsilon}{\sigma} \right),$$

where f^* and F^* are, respectively, the standard normal density and distribution functions, evaluated at $\lambda\epsilon/\sigma$, and $\sigma_*^2 = \sigma_u^2 \sigma_v^2 / \sigma^2$.

Thus, equations (7) and (9) provide estimates for u and v after replacing ϵ , σ_* , and λ by their estimates. Subtracting v from both sides of (7) results in

$$(10) \quad Q^* = f(X_a) - u = Q - v,$$

where Q^* is the firm's observed output adjusted

for the statistical noise captured by v . Equation (10) is used to compute the vector X_i and to algebraically derive the cost frontier which, in turn, is the basis for obtaining the minimum cost factor demand equations. The latter are then used to calculate the vector X_e .

Data and Empirical Model

Cross-sectional data for a sample of 511 New England dairy farms (excluding the state of Rhode Island) are used to estimate a Cobb-Douglas stochastic production frontier which is the basis for deriving a stochastic cost frontier and related efficiency measures. Following Zellner, Kmenta, and Drèze, we adopt the standard practice of justifying a single-equation model by assuming that farmers maximize expected profits (e.g., Caves and Barton, Kopp and Smith).

The data used in this study are from Dairy Herd Improvement (DHI) production records for the calendar year 1984 combined with data obtained from a mail survey. Descriptive statistics for the variables used in the analysis are given in table 1. The specific model estimated is the following:

$$(11) \quad \ln Q = \ln \beta_0 + \beta_1 \ln X_1 + \beta_2 \ln X_2 + \beta_3 \ln X_3 + \beta_4 D_1 + \beta_5 D_2 + \beta_6 D_3 + \epsilon,$$

where Q is annual milk production per farm measured in hundredweight adjusted to a 3.5% butterfat basis; X_1 is annual consumption of purchased dairy concentrate in tons per farm; X_2 is annual consumption of forage feed in tons per farm which is equal to the consumption of succulent roughage plus dry roughage, assuming a dry matter content of 30% and 90%, respectively; and X_3 is annual variable labor used per farm measured in full time worker-equivalents. The original data set contains information on total labor used per farm; thus, to obtain variable labor one operator per farm is assumed to be fixed, and this is subtracted from total labor. This is not an entirely adequate adjustment since family labor is often considered to be a fixed input (e.g., Stefanou and Saxena). However, data limitations preclude the construction of a more refined measure of labor. D_1 is a binary variable equal to one if the farm has a stanchion barn and a bucket and carry milking system and zero otherwise; D_2 is a binary variable equal to one if the farm is located in Maine or Massachusetts and zero otherwise; D_3 is a binary variable equal to one if the farm is located in Vermont and zero

Table 1. Ordinary Least Squares (OLS) Estimates of a Cobb-Douglas (C-D) Production Function and Maximum Likelihood (ML) Estimates of a C-D Production Frontier

Variable	Mean (std. dev.)	OLS Production Function Estimates (std. error)	ML Production Frontier Estimates (Asymptotic std. error)
Intercept		4.471*** ^b (.079)	4.557*** (.089)
C. Feed (X_1)	176.70 (143)	0.562*** (.019)	0.560*** (.019)
Forage (X_2)	230.38 (164)	0.340*** (.019)	0.343*** (.020)
Labor (X_3)	1.03 (0.94)	0.014*** (.004)	0.014*** (.005)
D_1 (%)	0.05	-0.081*** (.031)	-0.083*** (.037)
D_2 (%)	0.31	-0.063*** (.019)	0.060*** (.019)
D_3 (%)	0.50	0.063*** (.018)	0.064*** (.018)
Function coefficient		0.916	0.917
F-statistic model		1169***	
F-statistic CRTS ^a		43***	
Adj. R^2		0.932	
λ			0.779*** (0.369)
σ^2			0.167*** (.015)
Log likelihood			260.1

Note: This 1984 sample included 511 New England dairy farms.

^a CRTS constant returns to size.

^b Three asterisks indicate significance at 1% level; two indicate significance at 5% level.

otherwise; β_i are parameters to be estimated ($i = 0, \dots, 6$); and ϵ is a composed error term defined earlier. The dummy variables capture any effects of milking technology (D_1) and farm location (D_2 and D_3) on the placement of the production frontier.⁴ The variable inputs included in the specification of equation (11) are similar to those recently used by Stefanou and Saxena in their model of Pennsylvania dairy farms.

Results and Analysis

Table 1 presents the maximum likelihood estimates of the stochastic production frontier. For comparison purposes OLS estimates are also shown. All parameter estimates are statistically significant at the 1% level in both models. The only exception is the parameter for D_1 in the production frontier equation which is significant

at the 5% level. Slope parameters across equations are similar, which suggests that the frontier function represents a neutral upward shift of the OLS model. The quasi-function coefficient in both the frontier and the OLS model is about 0.92, and the hypothesis of constant returns to size, based on a restricted least squares regression, is rejected at the 1% level.

The dual cost frontier, derived analytically from the production frontier shown in table 1, is equal to

$$(12) \quad \ln C = 0.014 + 0.091D_1 + 0.660D_2 - 0.069D_3 + 0.611\ln P_{1i} + 0.374\ln P_{2i} + 0.016\ln P_{3i} + 1.091\ln Q^*$$

where C is variable cost of milk production per farm in 1984; D_j is a dummy variable ($j = 1, 2, 3$) as defined above; P_{1i} is price per ton of 16% protein dairy feed in the i th state ($i = 1, 2, \dots, 5$); P_{2i} is price per ton of forage feed in the i th state; P_{3i} is annual wages per worker employed in agriculture, forestry, and fishing in the i th state; and Q^* is annual milk production per

⁴ Connecticut and New Hampshire are combined because preliminary estimations showed that the location effect was almost identical for these two states. Maine and Massachusetts are also combined for the same reason.

Table 2. Frequency Distribution of Economic, Technical, and Allocative Efficiency Estimates

Efficiency Level (%)	Economic Efficiency	Technical Efficiency	Allocative Efficiency
>95	0 ^a (0.0)	0 (0.0)	94 (18.4)
>90 ≤95	0 (0.0)	0 (0.0)	57 (11.1)
>85 ≤90	1 (0.2)	70 (13.7)	102 (20.0)
>80 ≤85	70 (13.7)	407 (79.6)	108 (21.1)
>75 ≤80	77 (15.1)	31 (6.1)	73 (14.3)
>70 ≤75	112 (21.9)	3 (0.6)	38 (7.4)
>65 ≤70	127 (24.9)	0 (0.0)	25 (4.9)
>60 ≤65	65 (12.7)	0 (1.6)	8 (1.6)
≤60	59 (11.5)	0 (0.0)	6 (1.2)
Mean	70.2	83.0	84.6
Minimum	44.6	72.6	53.5
Maximum	86.3	87.7	99.7

Note: See note, table 1.

^a The top figure is the number of farms, and the figure in parenthesis is the percent of farms.

farm adjusted for any statistical noise as specified in equation (10) above. Input price data for individual farms are not available, which forces us to employ state-level average input prices for 1984.⁵ In general, this is a reasonable assumption considering the relatively small geographical area where the farms in the sample are located and the competitive nature of the markets in question.

The average economic (*EE*), technical (*TE*) and allocative efficiency (*AE*) indexes computed for the sample, shown in table 2, are 70.2, 83.0, and 84.6, respectively. This average technical efficiency measure of 83.0% is very close to the 82.0% reported by Bravo-Ureta based on a probabilistic production frontier model for a sample of New England dairy farms participating in ELFAC (Electronic Farm Accounts). However, the measure is considerably higher than the 69.0% reported by Tauer and Belbase, who

used a corrected ordinary least squares procedure to estimate a production frontier for a sample of New York dairy farms. Similar comparisons for economic and allocative efficiency cannot be made because, as far as we know, no such measures for dairy farms have been reported in the literature.

Table 3 presents a summary of observed, economic, and technically efficient costs along with potential savings for three groups each containing 10% (51 farms) of the farms with the lowest, middle, and highest level of economic efficiency. Based on analysis of variance (ANOVA) and the Kruskal-Wallis (K-W) test, the results indicate that these costs are significantly different across the three groups at the 1% level. The data suggest that the most economically inefficient group in the sample could save an average of \$1.44 per hundredweight if it became technically efficient and \$3.60 per hundredweight if it became economically efficient. In contrast, the figures for the group with the highest economic efficiency are \$0.87 and \$0.96, respectively.

The ANOVA and K-W tests are also used to evaluate the relationship between efficiency and four socioeconomic characteristics of the farms in the sample. The four characteristics analyzed are (a) farm size, measured by the average number of dairy cows in 1984; (b) education, measured by the number of years of schooling completed by the farm operator; (c) extension, measured by the number of extension meetings attended plus farm visits by extension agents during 1982 and 1983; and (d) experience, measured by the age of the farm operator. The results of the K-W and ANOVA tests, shown in table 4, reveal a statistically significant positive relationship between *TE* and farm size, while the opposite is true for *AE* and *EE*. This same pattern is observed for extension. By comparison, experience shows no significant association with *TE* but is inversely related to *EE* and *AE*. Finally, education appears to have no statistically significant association with efficiency in this sample. Even in the cases where a statistically significant difference is observed, the actual efficiency levels across the three farm groups are remarkably close to each other.

Our results concerning farm size and *TE* are consistent with the estimates reported by Grisley and Mascarenhas for Pennsylvania dairy farmers, by Tauer and Belbase for New York dairy farms, and by Bagi for mixed crop and livestock farms in Tennessee. In contrast, no connection between *TE* and farm size was reported by Bravo-Ureta for New England dairy farms, by Byrnes

⁵ The input prices used are (U.S. Bureau of the Census, U.S. Department of Agriculture):

Connecticut: $P_1 = \$172$; $P_2 = \$96.4$; $P_3 = \$11,842$.
 Maine: $P_1 = \$178$; $P_2 = \$79.7$; $P_3 = \$10,735$.
 Massachusetts: $P_1 = \$176$; $P_2 = \$94.6$; $P_3 = \$14,327$.
 New Hampshire: $P_1 = \$192$; $P_2 = \$96.1$; $P_3 = \$11,908$.
 Vermont: $P_1 = \$187$; $P_2 = \$83.3$; $P_3 = \$11,167$.

Table 3. Summary of Observed and Efficient Costs and of Potential Savings

Variable	Economic Efficiency ^a			Analysis of Variance F-Value	Kruskal- Wallis Test
	Lowest 10%	Middle 10%	Highest 10%		
Per hundredweight	(\$)				
Observed cost	8.47	6.95	5.55	256*** ^b	126***
Economically efficient cost	4.87	4.68	4.59	8***	15***
Technically efficient cost	7.03	5.73	4.68	229***	125***
Savings if economic efficient	3.60	2.27	0.96		
Savings if technical efficient	1.44	1.22	0.87		

Note: See note, table 1.

^a The efficiency comparisons are performed for three groups of farms defined as 10% of the farms with the lowest, the middle, and the highest average economic efficiency.

^b Three asterisks indicate significance at the 1% level.

et al. for Illinois grain farms, and by Bagi for crop farms in Tennessee. The lack of relationship observed in our sample between education and *TE*, and between experience and *TE*, although surprising, is in agreement with the findings of Tauer and Belbase for New York dairy farms. Our results regarding the relationships between the socioeconomic variables and *AE* and *EE* as well as between extension and efficiency cannot be compared to other U.S. samples because no comparable studies have been published. However, studies using data from developing countries have also yielded conflicting results concerning the relationship between *TE* and socioeconomic variables similar to those in-

cluded in this paper (e.g., Belbase and Grabowski, Huang and Bagi, Kalirajan).

Concluding Comments

This paper extends Kopp and Diewert's efficiency decomposition methodology from a deterministic to a stochastic framework and uses the latter to analyze efficiency in dairy production. The methodology developed here yields efficiency measures that are not distorted by statistical noise. In addition, the methodology makes it possible to calculate not only technical but also allocative and economic efficiency relying solely

Table 4. Statistical Tests of the Association between Average Efficiency Indexes and Four Socioeconomic Characteristics

Group ^a	Farm Size			Education			Extension			Experience		
	EE ^b	TE	AE	EE	TE	AE	EE	TE	AE	EE	TE	AE
1	72.0	82.6	87.2	70.8	82.8	85.6	71.2	82.7	86.1	71.7	82.8	86.6
2	69.0	82.7	83.5	70.2	82.9	84.7	70.3	82.8	84.9	70.2	83.1	84.5
3	69.3	83.5	83.1	69.9	83.0	84.3	69.2	83.3	83.2	68.7	83.0	82.8
<i>Kruskal-Wallis Test</i>												
Calc. KW	18.0	21.0	22.0	0.7	1.3	0.5	7.4	8.7	10.2	14.7	0.9	15.9
Significance	*** ^c	***	***				**	**	***	***		***
<i>Analysis of Variance</i>												
Calc. F	7.7	8.8	10.0	0.3	0.4	0.5	2.5	4.0	3.8	6.5	0.6	6.9
Significance	***	***	***				**	**	***	***		***

Note: See note, table 1.

^a The groups for each of the four socioeconomic characteristics are defined as follows:

Farm Size -	Group 1: <44 cows;	Group 2: 44-64 cows;	Group 3: >64 cows;
Education -	Group 1: <12 years of schooling;	Group 2: 12 years of schooling;	Group 3: >12 years of schooling;
Extension -	Group 1: no extension contacts;	Group 2: 1-7 extension contacts;	Group 3: >7 extension contacts;
Experience -	Group 1: <37 years of age;	Group 2: 37-52 years of age;	Group 3: >52 years of age.

^b The efficiency index is the average for each group.

^c Three asterisks indicate significance at 1% level; two, at the 5% level.

on the econometric estimation of a production frontier.

The analysis shows that, for our sample of New England dairy farms, average technical efficiency is 83.0%, average economic efficiency is 70.2%, and average allocative efficiency is 84.6%. These results suggest little difference, on average, between technical and allocative inefficiency, which is of interest given the controversy regarding the relative importance of the various components of efficiency (Leibenstein, 1966, 1978; Comanor and Leibenstein; Stigler). However, the analysis shows that focusing only on technical efficiency understates the benefits that could be derived by individual farmers as well as society from an improvement in overall performance. Analyses of the relationship between efficiency and four socioeconomic variables—farm size, education, extension, and experience—reveal that, despite some statistically significant associations, efficiency levels are not markedly affected by these variables. Further experimentation with different data sets and alternative methodologies is needed before the sensitivity of the results to the research design can be thoroughly judged.

[Received December 1988; final revision received July 1990.]

References

- Afriat, S. N. "Efficiency Estimation of Production Functions." *Int. Econ. Rev.* 13(1972):568–98.
- Aigner, D. J., and S. Chu. "On Estimating the Industry Production Function." *Amer. Econ. Rev.* 58(1968):826–39.
- Aigner, D. J., C. A. K. Lovell, and P. J. Schmidt. "Formulation and Estimation of Stochastic Frontier Production Function Models." *J. Econometrics* 6(1977):21–37.
- Akridge, J. T. "Measuring Productive Efficiency in Multiple Product Agribusiness Firms: A Dual Approach." *Amer. J. Agr. Econ.* 71(1989):116–25.
- Bagi, F. S. "Relationship Between Farm Size and Technical Efficiency in West Tennessee Agriculture." *S. J. Agr. Econ.* 14(1982):139–43.
- Belbase, K., and R. Grabowski. "Technical Efficiency in Nepalese Agriculture." *J. Develop. Areas* 19(1985):515–25.
- Bravo-Ureta, B. E. "Technical Efficiency Measures for Dairy Farms Based on a Probabilistic Frontier Function Model." *Can. J. Agr. Econ.* 34(1986):399–415.
- Byrnes, P., R. Färe, S. Grosskopf, and S. Kraft. "Technical Efficiency and Size: The Case of Illinois Grain Farms." *Eur. Rev. Agr. Econ.* 14(1987):367–81.
- Caves, R. E., and D. R. Barton. *Technical Efficiency in U.S. Manufacturing Industries*. A Report to the U.S. Department of Commerce, Washington DC, 1988.
- Comanor, W. S., and H. Leibenstein. "Allocative Efficiency, X-Efficiency and the Measurement of Welfare Losses." *Economica* 36(1969):304–9.
- Farrell, M. J. "The Measurement of Production Efficiency." *J. Royal Statist. Soc., Series A* 120(1957):253–81.
- Forsund, F. R., C. A. K. Lovell, and P. Schmidt. "A Survey of Frontier Production Functions and of their Relationship to Efficiency Measurement." *J. Econometrics* 13(1980):5–25.
- Greene, W. H. "Maximum Likelihood Estimation of Econometric Frontier Functions." *J. Econometrics* 13(1980):27–56.
- Grisley, W., and J. Mascarenhas. "Operating Cost Efficiency on Pennsylvania Dairy Farms." *Northeast. J. Agr. and Resour. Econ.* 14(1985):88–95.
- Hall, B. F., and E. P. LeVein. "Farm Size and Economic Efficiency: The Case of California." *Amer. J. Agr. Econ.* 60(1978):589–600.
- Huang, C. J., and F. S. Bagi. "Technical Efficiency on Individual Farms in Northwest India." *S. Econ. J.* 51(1984):108–15.
- Jondrow, J., C. A. K. Lovell, I. S. Materov, and P. Schmidt. "On the Estimation of Technical Inefficiency in the Stochastic Frontier Production Function Model." *J. Econometrics* 19(1982):233–38.
- Kalirajan, K. "Farm-Specific Technical Efficiencies and Development Policies." *J. Econ. Stud.* 11(1984):3–13.
- Kopp, R. J., and W. E. Diewert. "The Decomposition of Frontier Cost Function Deviations into Measures of Technical and Allocative Efficiency." *J. Econometrics* 19(1982):319–31.
- Kopp, R. J., and V. Kerry Smith. "Frontier Production Function Estimates for Steam Electric Generation: A Comparative Analysis." *S. Econ. J.* 47(1980):1049–59.
- Kumbhakar, S. C. "The Specification of Technical and Allocative Inefficiency in Stochastic Production and Profit Frontiers." *J. Econometrics* 34(1987):335–48.
- Leibenstein, H. "Allocative Efficiency vs. 'X-Efficiency'." *Amer. Econ. Rev.* 56(1966):392–415.
- . "X-Inefficiency Xists—Reply to an Xorcist." *Amer. Econ. Rev.* 68(1978):203–11.
- Meeusen, W., and J. van den Broeck. "Efficiency Estimation From Cobb-Douglas Production Functions with Composed Error." *Int. Econ. Rev.* 18(1977):435–44.
- Quiggin, J., and A. Bui-Lan. "The Use of Cross-Sectional Estimates of Profit Functions for Tests of Relative Efficiency: A Critical Review." *Aust. J. Agr. Econ.* 28(1984):44–55.
- Quiroga, R., and J. Elterich. *The Measurement of Technical Efficiency: Frontier Production Function Methodology Applied to a Sample of Dairy Farms in Pennsylvania*. University of Delaware, Agr. Exp. Sta. Bull. No. 466, 1986.
- Richmond, J. "Estimating the Efficiency of Production." *Int. Econ. Rev.* 15(1974):515–21.
- Schmidt, P. "Frontier Production Functions." *Econometric Reviews* 4(1985–86):289–328.
- . "On the Statistical Estimation of Parametric Frontier

- Production Functions." *Rev. Econ. and Statist.* 58(1976):238-39.
- Schmidt, P., and C. A. K. Lovell. "Estimating Technical and Allocative Inefficiency Relative to Stochastic Production and Cost Frontiers." *J. Econometrics* 9(1979):343-66.
- Stefanou, S. E., and S. Saxena. "Education, Experience, and Allocative Efficiency: A Dual Approach." *Amer. J. Agr. Econ.* 70(1988):338-45.
- Stigler, C. J. "The Xistence of X-Efficiency." *Amer. Econ. Rev.* 66(1976):213-26.
- Tauer, L. W., and K. P. Belbase. "Technical Efficiency of New York Dairy Farms." *Northeast. J. Agr. and Resour. Econ.* 16(1987):10-16.
- Taylor, T. G., H. E. Drummond, and A. T. Gomes. "Agricultural Credit Programs and Production Efficiency: An Analysis of Traditional Farming in Southeastern Minas Gerais, Brazil." *Amer. J. Agr. Econ.* 68(1986):110-19.
- Timmer, C. P. "On Measuring Technical Efficiency." *Food Res. Inst. Stud.* 9(1970):98-171.
- U.S. Bureau of the Census. *Statistical Abstract of the United States: 1987*. Washington DC, 1987.
- U.S. Department of Agriculture. *Agricultural Statistics: 1985*. Washington DC, 1985.
- Zellner, A., J. Kmenta, and J. Drèze. "Specification and Estimation of Cobb-Douglas Production Function Models." *Econometrica* 34(1966):784-95.

Prospect Theory and Risk Preferences of Oregon Seed Producers

Alan Collins, Wesley N. Musser, and Robert Mason

Prospect theory relates risk preference classifications to gains and losses from a reference income level. This study applies prospect theory to reinterpret historical studies of risk preferences of Oregon grass seed growers. A significant relationship between changes in classifications of preferences and changes in income was found. Results indicated that those who lost income were concentrated in the category of changing to risk preferrers. Income changes calculated from crop combinations were also found to be correlated in a theoretically correct pattern with positive measures of risk from crop enterprises. The research therefore is consistent with further applications of prospect theory to farm management.

Key words: expected utility theory, prospect theory, risk preferences, variance estimation.

Substantially different risk preferences have been reported among a single population of Oregon grass seed producers at different points in time (Halter and Mason, Whittaker and Winter). Similar intertemporal instability of risk preferences among agricultural producers has been reported in India (Binswanger). From the results in Oregon, Fleisher and Robison concluded that the elicitation methods were fundamentally flawed, or risk preferences do change over time.

In the Oregon research, the possibility of errors was minimized by survey research center data collection and polynomial curve fitting of utility functions. These methods minimize interview errors (Musser and Musser, Fleischer and Robison) and functional form specification errors (Buccola and French, Dillon and Scandizzo, Musser et al.), respectively. However,

measurement errors in variables could bias utility function coefficients (Knowles). In addition, response measurement errors may result if the elicitation process is beyond human cognitive capacity. In this case, changes in measured risk preferences could be the result of random errors in responses (Musser and Musser).

While methods problems cannot be dismissed, the alternative explanations of systematic changes in risk preferences do merit attention. Earlier research on these data indicated that changes in wealth could not explain the preference shifts within the expected utility framework. Another possibility is prospect theory, which predicts individuals can alternate between risk-averse and risk-seeking behavior (Kahneman and Tversky). This theory has been applied in natural resource economics to reconcile differences between willingness to pay and willingness to sell (Gregory; Cummings, Brookshire, and Schulze).

This paper investigates the use of prospect theory to rationalize changing risk preference results from the earlier Oregon studies (Halter and Mason, Whittaker and Winter). The next section of the paper summarizes relevant features of prospect theory and its implications for these earlier studies. The subsequent two sections present further empirical analysis of the risk preference data from these studies. Changes in risk preferences are related to changes in gross

The authors are, respectively, an assistant professor, Division of Resource Management, West Virginia University; a professor of agricultural economics, Pennsylvania State University, and a professor of sociology, Survey Research Center, Oregon State University. Collins was a graduate research assistant and Musser a visiting associate professor, Department of Agricultural and Resource Economics, Oregon State University, when this research was instituted.

Technical Paper No. 9267 of the Oregon Agricultural Experiment Station.

The late A. N. Halter participated in collecting the sample data and estimated utility functions. Brian Freeze, Jim Cornelius, and Steve Buccola made helpful suggestions on the research in this paper, and William McSweeney and two anonymous referees made helpful comments on earlier drafts of this paper.

farm income as a measure of wealth changes between sample periods. In addition, managerial responses in enterprise organization are compared to gross income changes to link wealth changes with managerial behavior.

Aspects of Prospect Theory

While prospect theory has roots in expected utility theory, it is strongly influenced by cognitive theories of bounded rationality (Simon). Kahneman and Tversky developed the theory to rationalize divergences in behavior from predictions of expected utility theory. Their earlier work on decision heuristics (Tversky and Kahneman) strongly influenced this theory. Prospect theory describes decisions as having two stages—editing and evaluation. Editing simplifies the decision problem for the evaluation stage. For example, alternative outcomes of a decision are redefined as gains and losses from a reference point (P_r), which is usually the current wealth or equity capital position of an individual.

Evaluation utilizes an adaption of the expected utility decision rule:

$$(1) \quad V(D_j) = \sum_{i=1}^n \pi(p_i) v(x_i),$$

where D_j is a decision alternative, x_i is a gain or loss outcome with a probability p_i , n is the number of edited outcomes, $\pi(p_i)$ is a weighting function based on p_i , and $v(x_i)$ is a value function. The weighting function is related to subjective probabilities but does not exactly follow probability rules. Specifically, $\pi(p_i) > p_i$ for small p_i and $\sum_{i=1}^n \pi(p_i) < 1$. In evaluation, the alternative from the edited prospects that maximizes $V(D_j)$ is chosen.¹

The value function has a similar role to a utility function but is defined on gains and losses in wealth rather than level of wealth. As illustrated in figure 1, specific properties are pos-

tulated for $v(x_i)$. For $x_i > 0$, the individual is risk averse, $v''(x) < 0$, and, for $x_i < 0$, an individual is a risk seeker, $v''(x_i) > 0$. A discontinuity in $v(x)$ is assumed at $x_i = 0$ with $v'(x_i) < v'(-x_i)$ for $x_i > 0$. Kahneman and Tversky proposed the properties of $v(x_i)$ based on experimental evidence of individual probabilistic decisions. Schurr recently confirmed this shape in experiments on risky business negotiations on commodities.

To predict changes in risk preferences and behavior over time with prospect theory, the concept P_r is used. Prediction is possible in situations where decisions are based on a P_r which differs from the status quo in individual wealth. These situations can occur when decision makers have not yet adapted to recent changes in wealth or when expectations of future wealth positions have been incorporated into P_r . In such cases, gains and losses from a decision are evaluated in connection with recent wealth changes or relative to an expected level of wealth.² The effect of a failure to incorporate losses into P_r is illustrated in figure 1. Assume a recent wealth loss (Y_L), which is not incorporated into P_r , so that $v(x_i)$ is stable. Then, prospects are evaluated as $v(Y_L + x_i)$. Now, even gains are evaluated with a convex value function as long as Y_L

² Admittedly this failure to adapt P_r is ad hoc. However, such ad hoc dynamics are also used in economics research. For example, Deaton and Muehlbauer provide a review of such inflexibility in consumption; Langemeier and Patrick is a recent application of such models to consumption behavior of farm families.

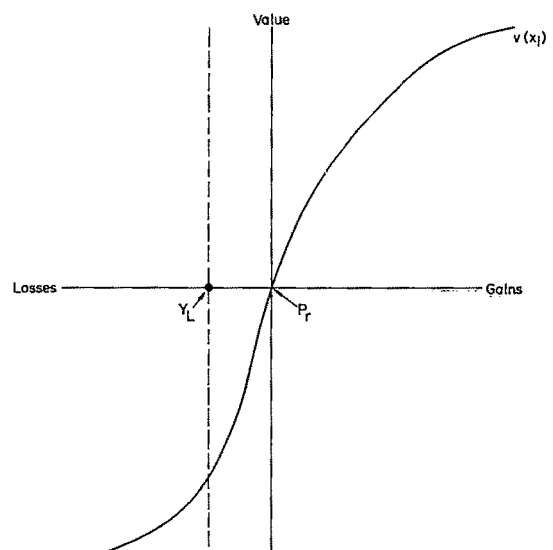


Figure 1. Value function hypothesized from prospect theory

¹ Prospect theory, like most psychological theories, is a positive rather than normative theory. Unlike decision theory, all the theoretical constructs need not be empirically estimated. The focus is on their behavioral implications rather than postulates. Positive economics is similar in that respect; consumer utility functions and firm production functions need not be estimated to test theoretical hypotheses. For example, Kahneman and Tversky do not expect $v(x_i)$ and $\pi(p_i)$ to be estimated. Nevertheless, some empirical findings in farm management support their assumptions. Kahneman and Tversky cited research on utility functions from agricultural managers as being consistent with $v(x_i)$ as do Anderson, Dillon, and Hardaker. Bessler and SriRamaratnam et al. report subjective probability distributions that have some properties similar to $\pi(p_i)$.

+ $x_i < 0$. Thus, individuals in a loss situation would exhibit risk-seeking behavior. This failure to adapt to recent wealth changes for individuals in a loss situation can be related to commonly perceived behavior. Kahneman and Tversky cite a well-known tendency toward betting on "long shots" near the end of a betting day as an example of failure to adapt P_r . A failure to adapt to gains is harder to relate to common behavior (Kahneman and Tversky, pp. 286–88). Logically, this failure would have the opposite effect of increasing risk aversion. Isen, Nygren, and Ashby recently found increasing risk aversion in individuals that received a prior gain. However, they explained their results as a shift in $\pi(p_i)$ rather than P_r .

In this study, Pratt's absolute risk aversion coefficient (R) for individual producers is used to measure responses of P_r to gains and losses in wealth. Prospect theory provides no prediction of the changes in R if changes in wealth are incorporated in P_r . However, assuming failure to incorporate prior wealth changes in P_r gives definite predictions for changes in R . Recent losses (gains) in wealth result in $v(x_i)$ being convex (concave) for both gains and losses. Thus, measures of R would tend to be negative (positive) after the wealth change. Whatever the measure of R before the wealth changes, intertemporal changes in R therefore should tend to have the same sign as recent changes in wealth. This implication of prospect theory is evaluated for the Oregon grass seed producers in this research.

Methods

Data for this study are from three surveys of Oregon grass seed producers in March 1973, April 1974, and December 1975. Data on crop acreages and other farm characteristics were collected for the 1973, 1974, and 1975 crop years. Risk preferences were elicited in the last two surveys based on gross farm income from completed 1973 and 1974 crop years, respectively. Thus, cropping decisions in the last two surveys lagged realized income on which risk preferences were based. Of the forty-four producers interviewed in 1974, thirty-seven completed the 1975 survey.

Risk Preferences

Risk preferences were elicited with the Ramsey method (Halter and Mason). Gross farm income

was used for accuracy in recall during the interviews. Net income (or wealth) would have been more desirable theoretically. However, gross income does correspond with the standard assumption of the stochastic component of net income in production (Dillon). Individual utility functions were fit to linear, quadratic, or cubic functional forms based on goodness of fit criteria and visual inspection of scatter plots (Halter and Mason). These criteria allowed functional form to vary from 1974 to 1975. Absolute risk aversion coefficients (R) were calculated from the estimated functions at sample-reported 1973 and 1974 gross farm income levels.

Change in gross farm income between 1973 and 1974 ($\Delta_I = 1974 \text{ income} - 1973 \text{ income}$) was used to measure wealth changes. More accurate measures of wealth changes would include data on changes in net farm income, equity capital, household consumption spending, and nonfarm income. These data are harder to collect accurately and were not included in the surveys of Oregon grass seed producers. With this wealth change measure between 1973 and 1974, the absolute risk aversion coefficient prior to the wealth change was compared to that after the change ($\Delta_R = R_{1974} - R_{1973}$). A chi-square classification table was used to test for independence between Δ_I and Δ_R as measures of wealth and risk preference changes, respectively.

Behavioral Risk Responses

Given the behavioral focus of prospect theory, this study also considers the managerial response in crop enterprise organization among the Oregon grass seed producers to the changes in risk preferences. While measurement of risk levels of decisions is a difficult theoretical and empirical issue (Hey, Young), this study assumed variance of gross income (VAR_I) approximated risk in behavior (Levy and Markowitz, Robison and Barry). While subjective probabilities would have been more consistent with decision theory and reflected differences in land, management, and other resources among the farmers, they were unavailable for this study. Thus, historical risk measures were used to proxy individual subjective distributions.

Crops included in this analysis were annual rye, perennial rye, tall fescue, bentgrass, Merion bluegrass, wheat, oats, and barley. Some producers had livestock and/or other crops; however, inadequate sample data were available to include these enterprises in variance of gross

income. A time series of annual acreage, price, and yield data for each of the crops were collected for 1954 to 1974 (Oregon State University Extension Service). County averages for Linn County, where the majority of the sample producers were located, were used for the analysis. Gross incomes were inflated to 1974 price levels with the "Index of Prices Paid by Farmers for Production Items" (U.S. Department of Agriculture).

Using criteria of Peck and Young for positive objective risk measures, risk measures ($V(X, Y)$) were calculated as weighted moving averages of historical forecast errors of gross income:

$$(2) \quad V(X, Y) = \sum_{t=1}^n w_t (X_t - E_t(X)) (Y_t - E_t(Y)),$$

where $E_t(X)$ and $E_t(Y)$ are forecasted expected values of X and Y in time period t , and w_t are weights on forecast errors from previous periods where $\sum_{t=1}^n w_t = 1$. $V(X, Y)$ is a mean-squared error estimate of variance of a variable when $X = Y$ and of covariance between X and Y when $X \neq Y$.

This research utilized an adaptive expectation approach to jointly specify forecasted expected values and weights. Following Nerlove, expected values, $E_t(Z)$ were calculated with estimated coefficients of expectations (b_z) from supply response models:

$$(3) \quad E_t(Z) = b_z Z_{t-1} + (1 - b_z) Z_{t-2} + (1 - b_z)^2 Z_{t-3} + \dots$$

Because supply response models including risk variables (Just) were unavailable, separate weights could not be used on expected values and risk indices. Thus, b_z coefficients were also used as estimates of w_t . Published estimates of b_z were adapted for this research. The average of estimates for coefficients of expectation (0.21) for similar crops in the Willamette Valley (Ryan, Conklin, and Edwards) were used for grass seed crops. Similar models were unavailable for grains in Oregon or the Pacific Northwest. Therefore, a value of 0.77 from a national model for this period (Lidman and Bawden) was used for grains because similar models were unavailable for Oregon or the Pacific Northwest. Minor adjustments were made in the geometric patterns of weights to account for data availability and to insure that they summed to one (table 1). A simple average of the two weights was used to es-

Table 1. Weightings Used for Expected Values of Crop Gross Incomes and Mean Forecast Errors

Year	Grass Seed Crops	Grain Crops	Grass-Grain Covariance
1	0.232	0.75	0.491
2	0.183	0.20	0.192
3	0.145	0.05	0.097
4	0.114		0.057
5	0.090		0.045
6	0.071		0.036
7	0.056		0.028
8	0.045		0.022
9	0.035		0.018
10	0.028		0.014

timate covariances between grass seeds and grains.³

Separate variance-covariance matrices for crop gross incomes were calculated for 1974 and 1975. These estimated risk matrices were appropriate for the 1974 and 1975 planting expectations, respectively. Estimates of VAR_t for both crop portfolios were then formed with individual producer crop acreages and the estimated risk matrices.

These variances of gross farm income were used to test the behavioral implications of prospect theory arising from wealth changes. Forecasted gross income variance for 1974 crop selection was the behavioral measure prior to the wealth change; this same variance for 1975 crop selection was used for subsequent behavior measurement. Changes in variance ($\Delta_v = VAR_{1975} - VAR_{1974}$) were negative for twenty-eight of the thirty-seven observations. The large increase in realized grain prices in 1973 resulted in a large forecast error which was weighted by 0.75 in 1974. Grain prices stabilized in 1974 so the most recent forecast error was much smaller for 1975, and VAR_{1975} was smaller than VAR_{1974} for many producers. With such a large proportion nega-

³ A potential weakness of this method of forming $E_t(Z)$ and w_t is the fundamental identification problem between adaptive expectations and partial adjustment models. Nerlove recognized this problem, and Doran recently provided some statistical methods to test for these alternative explanations in an estimation context. Obviously, the differential length of weights in table 1 could be consistent with partial adjustment behavior of annual versus perennial crops. With that interpretation, these weights are inappropriate for estimating statistical moments. On the other hand, it is plausible that a longer time period would be used in forming expectations on perennials simply because of their longer future production potential. Therefore, these estimates were assumed to reflect adaptive expectations so that an empirical source of w_t was available.

tive, a chi-square test with Δ_v would not be meaningful. Therefore, risk in cropping responses was evaluated with Spearman's rank correlation coefficients (ρ_s) between Δ_v and other variables. Student's t -tests were used to test $H_0: \rho_s = 0$ against $H_1: \rho_s > 0$ (Siegal).

Correlation analyses were also used to relate estimated gross income to Δ_v . A large number of the sample reported the same gross increase in 1973 and 1974 so that Δ_i had a concentration of zeroes. Estimated gross incomes realized in 1973 and 1974 were therefore substituted for sample estimates. Changes in estimated gross income ($\Delta_G = 1974$ gross income–1973 gross income) and percentage change in estimated gross income between 1973 and 1974 ($\%\Delta_G$) were calculated with realized county average gross income per acre and sample crop acreages. Because prior losses in wealth are hypothesized to lead to increasing risky behavior, a negative relationship is expected between Δ_v and Δ_G or $\%\Delta_G$. In the correlation analysis, $-\Delta_G$ and $-\%\Delta_G$ were used so that hypothesized coefficients would be positive.

Results

Tabulations of risk preference classifications for the 1974 and 1975 surveys are reported in table 2. These data demonstrate the instability between the two samples reported in Whittaker and Winter. The only stability was in total numbers in the risk-preferring category—eleven in 1974 and twelve in 1975. Especially surprising is the small number of producers with the same classification in both years, only nine out of thirty-seven. Not surprisingly, $X^2_{(4)} = 4.09$ so that the hypothesis that the samples are independent cannot be rejected.

The relationship of Δ_R to Δ_i is presented in table 3. More than 40% of the sample reported the same gross income. Both the same and more

Table 3. Changes in Gross Farm Income and Risk Preferences for Oregon Grass Seed Producers, 1974 to 1975

Absolute Risk-Averse Coefficient Change (Δ_R)	Gross Income Change (Δ_i)			
	Less	Same	More	Total
Risk seeker ^a	5	7	4	16
Risk avoider ^b	1	7	4	12
Same	0	2	7	9
Total	6	16	15	37

^a Coefficient changes of: (a) positive to zero or negative, and (b) zero to negative.

^b Coefficient changes of: (a) negative to zero or positive, and (b) zero to positive.

income categories had equal numbers moving toward risk-seeking and risk-avoiding preferences. The ambiguity in the more income category is related to an unknown pattern of adapting their reference point to increases of wealth. Isen, Nygren, and Ashley also had ambiguous results for income gains. Individuals with less gross farm income and consequently smaller wealth likely would not have adapted their P_i and, thus, evaluate gains as $V(Y_L + x_i)$ in figure 1. While the numbers are small, five out of six changed to a more risk-seeking classification. This finding is consistent with the implications of prospect theory. The chi-square test of independence supported this interpretation: $X^2_{(4)} = 9.89$ so that the null hypothesis of no association can be rejected at the 5% level of significance. Correlation analysis of Δ_G with Δ_R and Δ_v is reported in table 4. The sample correlation coefficient between $-\Delta_G$ and Δ_v equaled 0.30. With $t_{(35)} = 1.88$, $\hat{\rho}_s$ supports H_1 of $\rho_s > 0$ at

Table 4. Spearman's Rank Correlation Matrix of Behavioral Relationships Related to Prospect Theory

	Change in Gross Income ($-\Delta_G$)	Percentage Change in Gross Income ($-\%\Delta_G$)	Change in Absolute Risk Aversion ($-\Delta_R$)	Change in Variance of Gross Income (Δ_v)
$-\Delta_G$			0.12 (0.50) ^a	0.30 (1.88)
$-\%\Delta_G$			0.26 (1.60)	0.33 (2.07)
$-\Delta_R$				-0.06 (-0.36)

^a Student's t -statistics in parentheses.

Table 2. Classification of Risk Preferences, Oregon Grass Seed Producers, 1974 and 1975

Category	1974 Classification	1975 Classification		
		Preferring	Neutral	Averse
		(Number)		
Preferring	11	1	8	2
Neutral	14	6	6	2
Averse	12	5	5	2
Totals	37	12	19	6

the 5% level of significance. In addition, Δ_V and $-\% \Delta_G$ had a $\hat{\rho}_s$ of 0.33, with a $t_{(35)} = 2.072$. With this statistic, H_0 can be rejected at the 2.5% level. These tests showed that estimated gross income declines were correlated with increases in risk-seeking behavior.

The $\hat{\rho}_s$ value between $-\Delta_G$ and $-\Delta_R$ was only 0.12 and between $-\% \Delta_G$ and $-\Delta_R$ was 0.26. The signs of these statistics are consistent with the chi-square analysis which showed reduction in gross income leading to risk-seeking preferences. The latter was significant at the 10% level, which confirms the potential empirical relevance of $\% \Delta_G$ rather than Δ_G . Perhaps, small relative changes in wealth, which could have a large absolute value for large farms, are simplified to zero in the editing stage. The $\hat{\rho}_s$ between $-\Delta_R$ and Δ_V is quite small and insignificant. Δ_R likely had more measurement error than Δ_G .

The significant relationships between $-\Delta_G$ and Δ_V , and $-\% \Delta_G$ and Δ_V support the prediction of prospect theory used in this study that declines in wealth can result in increasingly risky behavior. This relationship does not necessarily imply that the sample tended to become risk seekers with declines in wealth but only that absolute risk aversion coefficients decreased as wealth decreased.

Conclusions

This paper reinterpreted the historical studies of risk preferences of grass seed producers in the Willamette Valley of Oregon from the perspective of prospect theory. Prospect theory was used to explain intertemporal instability of risk preferences through an assumption of a failure to incorporate prior changes in wealth into a decision maker's reference point for evaluation for outcome gains and losses. By using this assumption within the framework of prospect theory, changes in Pratt absolute risk aversion coefficients were significantly related to reported prior gross income changes in a manner consistent with prospect theory. In addition, a significant relationship was found between risk behavior as predicted by prospect theory and wealth changes measured with changes in gross income estimated from reported crop acreages.

This analysis is subject to numerous empirical limitations. Besides the well-known measurement errors in methods to estimate risk preferences and objective probability distributions, assumptions were included on measuring risk with

variances, using gross farm income as a wealth change measure, and only using the field crops to calculate gross farm income. Given these problems, the evidence supporting prospect theory is especially noteworthy. However, the results are not supportive of elicitation of risk preferences. Absolute risk aversion coefficients from elicited data did not predict risky behavior as well as changes in income.

This historical evidence supports further study of the application of prospect theory in farm management and finance. For example, the theory is an alternative to other explanations of behavior related to the recent farm financial stress problems such as Robison, Barry, and Burghardt and Featherstone et al. Prospect theory would predict that farmers in a loss situation would assume riskier managerial position. If financial risk increases due to accumulated unpaid debt, business risk would not be reduced, and may be increased. This pattern would suggest that farmers as a whole would have risk-preferring preferences consistent with a worsening loss situation. Future applications of prospect theory would require additional refinement of the relationships between prospect theory and the measures used in this paper. More important, such work should emphasize (a) direct behavioral linkages with changes in income and (b) the research and extension implications of this theoretical relationship.

[Received May 1989; final revision received July 1990.]

References

- Anderson, J. R., J. L. Dillon, and J. B. Hardaker. *Agricultural Decision Analysis*. Ames: Iowa State University Press, 1977.
- Bessler, D. "Forecasting in Risk Research." *Risk Analysis for Agricultural Production Firms: Concepts, Informational Requirements and Policy Issues*, pp. 95-106. Dep. Agr. Econ. Staff Pap. 85-85, Michigan State University, Nov. 1985.
- Binswanger, H. "Attitudes Toward Risk: Experimental Measurement in Rural India." *Amer. J. Agr. Econ.* 62(1980):395-407.
- Buccola, S., and B. French. "Estimating Experimental Utility Functions." *Agr. Econ. Res.* 30(1978):37-43.
- Cummings, R. G., D. S. Brookshire, and W. D. Schulze. *Valuing Environmental Goods: An Assessment of the Contingent Valuation Methods*. Totowa NJ: Rowman and Allanheld, 1986.
- Deaton, A. and J. Muehlbauer. *Economics and Consumer Behavior*. Cambridge: Cambridge University Press, 1980.
- Dillon, J. L. *The Analysis of Response in Crop and Live-*

- stock Production, 2nd ed. New York: Pergamon Press, 1977.
- Dillon, J. L., and P. Scandizzo. "Risk Attitudes of Subsistence Farmers in Northeast Brazil: A Sampling Approach." *Amer. J. Agr. Econ.* 60(1978):425-35.
- Doran, H. E. "Specification Tests for the Partial Adjustment and Adaptive Expectations Models." *Amer. J. Agr. Econ.* 70(1988):713-23.
- Featherstone, A. M., C. M. Moss, T. G. Baker, and P. V. Preckel. "The Theoretical Effects of Farm Policies on Optimal Leverage and the Probability of Equity Losses." *Amer. J. Agr. Econ.* 70(1988):572-79.
- Fleisher, B. and L. J. Robison. *Applications of Decision Theory and the Measurement of Attitudes Towards Risk in Farm Management Research in Industrialized and Third World Setting*. Dep. Agr. Econ., MSU Int. Develop. Pap. No. 6, Michigan State University, 1985.
- Halter, A. N., and R. Mason. "Utility Measurement for Those Who Need to Know." *West. J. Agr. Econ.* 3(1978):99-109.
- Hey, J. D. *Uncertainty in Microeconomics*. New York: New York University Press, 1979.
- Isen, A. M., T. E. Nygren, and F. G. Ashby. "Influence of Positive Affect on the Subjective Utility of Gains and Losses: It Is Just Not Worth the Risk." *J. Personality and Soc. Psych.* 55(1988):710-17.
- Just, R. E. "An Investigation of the Importance of Risk in Farmers' Decisions." *Amer. J. Agr. Econ.* 56(1974):14-25.
- Kahneman, D., and A. Tversky. "Prospect Theory: An Analysis of Decision Under Risk." *Econometrica* 47(1979):263-91.
- Knowles, G. J. "Some Econometric Problems in the Measurement of Utility." *Amer. J. Agr. Econ.* 66(1984):505-10.
- Langemeier, M. E., and G. F. Patrick. "Farmers' Marginal Propensity to Consume: An Application in Illinois Grain Farms." *Amer. J. Agr. Econ.* 72(1990):309-16.
- Levy, H., and H. M. Markowitz. "Approximating Expected Utility by a Function of Mean and Variance." *Amer. Econ. Rev.* 69(1979):308-17.
- Lidman, R., and D. Bawden. "The Impact of Government on Wheat Acreage." *Land Econ.* 50(1974):327-35.
- Musser, W. N., and L. M. Musser. "Psychological Perspectives on Risk Analysis." *Risk Management in Agriculture*, ed. P. J. Barry, pp. 82-94. Ames: Iowa State University Press, 1984.
- Musser, W. N., M. E. Wetzstein, S. Y. Reece, L. M. Musser, P. E. Varca, and C. C. J. Chou. "Classification of Risk Preferences with Elicited Utility Data: Does Functional Form Matter?" *West. J. Agr. Econ.* 9(1984):332-28.
- Nerlove, M. "Estimates of the Elasticities of Supply of Selected Agricultural Commodities." *J. Farm Econ.* 38(1956):496-509.
- Oregon State University Extension Service. Unpublished data of annual acreage, yield and price received by county in Oregon. Office of Economic Information, Dep. Agr. and Resour. Econ., Corvallis.
- Peck, A. E. "Hedging and Income Stability: Concepts Implications, and an Example." *Amer. J. Agr. Econ.* 75(1975):410-30.
- Robison, L. J., and P. J. Barry. *The Competitive Firm's Response to Risk*. New York: Macmillan Co., 1987.
- Robison, L. J., P. J. Barry, and W. G. Burghardt. "Borrowing Behavior Under Financial Stress by the Proprietary Firm: A Theoretical Analysis." *West. J. Agr. Econ.* 12(1987):144-51.
- Ryan, J. T., F. S. Conklin, and J. A. Edwards. *Demand and Supply in the Oregon Grass Seed Industry: An Economic Analysis*. Oregon State University Agr. Exp. Sta. Bull. No. 652, 1981.
- Schurr, P. H. "Effects of Gain and Loss Decision Frames on Risky Purchase Negotiations." *J. Appl. Psych.* 72(1987):351-58.
- Siegel, S. *Nonparametric Statistics for the Behavioral Sciences*. New York: McGraw-Hill Book Co., 1956.
- Simon, H. A. "Theories of Decision Making in Economics and Behavioral Science." *Amer. Econ. Rev.* 49(1958):253-83.
- SriRamaratnam, S., David A. Bessler, M. Edward Rister, J. E. Matocha, and James Novak. "Fertilization Under Uncertainty." *Amer. J. Agr. Econ.* 69(1987):349-57.
- Tversky, A., and D. Kahneman. "Judgment Under Uncertainty: Heuristics and Biases." *Science* 185(1974):1124-31.
- U.S. Department of Agriculture, National Agriculture Statistics Service. *Agricultural Prices, Annual Price Summary*. Washington DC, various issues.
- Whittaker, J. K., and J. R. Winter. "Risk Preferences of Farmers: An Empirical Example." *Risk Analysis in Agriculture; Research and Educational Development*. Dep. Agr. Econ. AE-4492, University of Illinois, 1980.
- Young, D. L. "Evaluating Procedures for Computing Objective Risk from Historical Time Series." *Risk Analysis in Agriculture: Research and Educational Development*. Dep. Agr. Econ. AE-4492, University of Illinois, 1980.

Robustness of the Mean-Variance Model with Truncated Probability Distributions

Steven D. Hanson and George W. Ladd

The known sufficient conditions for the mean-variance framework to produce expected utility results are violated in the presence of truncated probability distributions. A theoretical simulation is conducted to examine the ability of the linear mean-variance model to approximate expected utility results when the income distribution is truncated by the use of commodity option contracts. The mean-variance model is shown to produce solutions that are close approximations to the expected utility model results under the assumptions of constant absolute risk aversion and normally distributed prices. However, some inconsistency was found between the comparative static results of the two models.

Key words: futures, mean-variance, options, risk management.

The mean-variance (MV) framework has been a popular method for ordering choices into efficient and inefficient sets since its development by Markowitz. The problem of portfolio selection is reduced to choosing from a set of alternatives that provide minimum variance for given levels of expected return (i.e., the MV efficient set). The MV framework has been used to study a wide range of economic decisions made under risk such as the allocation of fixed assets in the presence of uncertain production processes (Freund), the demand for money (Tobin), and corporate financial decisions (Rubenstein). For a review of applications of mean-variance analysis in agricultural economics, see Robison and Brake; and Musser, Mapp, and Barry. Much of the previous research on risk management with futures contracts has employed the MV framework (e.g., see Anderson and Danthine, Kahl). Recently, commodity options contracts have become available as a risk management tool. Wolf has incorporated options contracts into an MV framework.

The inclusion of options in a decision maker's (DM's) portfolio results in truncated probability distributions that violate the standard assumptions used to justify the use of an MV representation of expected utility. Indeed, a variety of risky decision problems can be characterized by truncated probability distributions: bankruptcy, starvation, insurance schemes, limited liability, and price support and deficiency payment programs. This study examines the approximating ability of the standard linear MV model when the income probability distribution is truncated by the use of commodity options; it examines the robustness of the linear MV model to a violation of MV assumptions.

The Mean-Variance Model

The MV framework focuses on the first two moments of the underlying probability distribution. Its popularity can be traced to the tractable theoretical results it produces and to its computational convenience. Unfortunately, its straightforward nature results from some restrictive assumptions that are violated when options are included in a portfolio.

Specifying an expected utility function in terms of the first two moments of the underlying attribute's distribution has been shown to be consistent with the expected utility hypothesis only if at least one of the following sufficient con-

Steven D. Hanson is an assistant professor of agricultural economics at Michigan State University; George W. Ladd is a professor of economics and Charles F. Curtiss Distinguished Professor in Agriculture, Iowa State University.

Journal Paper No. J-13596 of the Iowa Agriculture and Home Economics Experiment Station, Project No. 2858. Partial funding support was provided by the Michigan State University Experiment Station.

The authors are grateful to anonymous referees for their valuable suggestions.

ditions is met: (a) the DM's utility function is quadratic (Tobin), (b) the DM has a concave utility function, and the random attribute is normally distributed (Samuelson), (c) the random attribute is a monotonic linear function of a single random variable (Meyer); sometimes referred to as the location-scale condition. Because these are only sufficient conditions, the MV framework may be consistent with the expected utility (EU) hypothesis under other conditions.

These sufficient conditions are often argued to be either violated or unacceptable. Baron shows that quadratic utility is a necessary and sufficient condition for MV representation of expected utility to be valid when the random variables are not restricted. Unfortunately, quadratic utility usually is viewed as an unacceptable preference function because it implies increasing absolute risk aversion and produces negative marginal utility as income rises above some given level.

Some have justified the use of the MV framework by assuming that condition (b) is satisfied (e.g., Anderson and Danthine). Others have criticized this assumption because few random variables take on values from negative to positive infinity and are symmetrically distributed, as is implied by the normality assumption.

Many utility functions are concave functions of income and can be used with a normally distributed income variable to yield an MV representation of the expected utility function. Only the negative exponential utility function, $U(y) = \alpha - \exp(-Ay)$, where $\alpha > 0$, in conjunction with a normally distributed attribute, y , will result in an expected utility function identical to the "standard" linear MV model, which is typically specified as¹

$$(1) \quad \text{Max}_x \hat{U} = \mu_y - (A/2) \sigma_y^2,$$

where μ_y is expected value of end-of-period income, A is the Arrow-Pratt measure of absolute risk aversion, σ_y^2 is variance of end-of-period income, and x is a vector of choice variables. The negative exponential utility function implies that the DM exhibits constant absolute risk aversion (CARA) and increasing relative risk aversion. The literature on risk preference characteristics

is somewhat inconclusive, although there is some evidence that decision makers exhibit decreasing absolute risk aversion (King and Robison). Regardless, most studies have been willing to accept the assumption of CARA to obtain the desirable properties of the standard MV model, although not all studies employing MV have used the negative exponential function.

Meyer and Robison have used condition (c) to justify the MV framework. Meyer has shown that the MV and EU frameworks are consistent regardless of the distribution of the random variable whenever a random outcome variable is a monotonic linear function of a single underlying random variable. Current work by Nelson is attempting to generalize Meyer's results to a class of models containing multivariate risks which are elliptically symmetric.

Some risky decision-making environments violate the known sufficient conditions for consistency. Newbery and Stiglitz recognize that risk is not symmetric for some decisions such as insurance schemes, and MV analysis may be seriously misleading for those decisions. One implication is that the MV framework may produce significant errors when the income distribution is truncated in the tails.

Much of the debate involving MV and EU models has focused on the approximating ability of the MV model. Levy and Markowitz show that the MV model provides reasonable approximations to EU maximization when the opportunity set contains a low probability of extreme outcomes. Tsiang shows that an MV framework approximates the EU maximization model when certain restrictions are placed on the skewness of the income distribution. Meyer and Rache show that, with "few enough" observations, there is no statistically significant difference between MV and EU rankings because the random alternatives are estimated imprecisely. Porter shows that MV sets of randomly constructed stock portfolios are consistent with EU models, except for portfolios having small means and variances. Tew and Reid explore the performance of the MV framework using firm specific, non-tradable investments for a representative farm in Georgia. Their results indicate the MV framework performs well in selecting investments that maximize expected utility. Robison and Barry argue for the use of the MV model on the basis of its ability to produce useful analytical results when more general models become too complex. Thus, there is some support, both empirical and deductive, for use of the MV framework even when its sufficient conditions are

¹ Robison and Barry formulate a linear MV model by claiming that A is the slope of a linear tangent at equilibrium to the isoutility line and its MV-efficient set. The assumptions of constant, increasing, and decreasing absolute risk aversion are then introduced through slope changes.

violated. In what follows, we show that conditions (b) and (c) are violated when the income distribution is truncated by including options in a portfolio. Then we examine the robustness of the linear MV model in the presence of truncated probability distributions for a DM with constant absolute risk aversion.

Income Probability Distribution with Commodity Options

To derive the distribution of income when options are available, assume that the DM has the amount Q of a single asset that may be sold at a specified future date in the cash market for some unknown spot price, P , or in the futures market for a certain price, P_f , or the DM may purchase the option to sell the output for a certain price, P_e . The output Q may be divided among the three markets in any way. The DM may sell a total amount in the cash, futures, and options markets that is greater than Q , may go long in the futures market, or write put options instead of selling them. There are assumed to be no transaction costs, basis uncertainty, or margin calls.

The only uncertainty faced by the DM is the output price P , and the DM can reduce uncertainty by marketing output in the futures or the options market. The DM's choice variables are the amounts marketed in the futures and options markets. The model is single period in nature and assumes that the initial positions are maintained until the end of the period at which time all cash, futures, and options transactions will take place. The quantity of output to be sold, the futures contract selling price, the exercise price, and the put option premium are known, exogenous variables to the DM. The DM's end-of-period income can be represented as equation (2):²

$$(2) \quad y(P) = PQ + x(P_f - P) + Iz(P_e - P) - zk,$$

where $y(P)$ is total end-of-period income, P is random end-of-period spot price, P_f is localized futures price for end-of-period delivery, Q is certain end-of-period output level, I is an indicator variable ($= 1$ if $P < P_e$, or $= 0$ if $P \geq$

P_e), P_e is localized exercise price of a put option, k is put option premium per bushel, x is the number of bushels of futures contracts, z is the number of bushels of put option contracts.

Expression (2) is derived from two distinct income functions:

$$(3) \quad y(P) = PQ + x(P_f - P) + z(P_e - P) - zk \quad \text{if } P < P_e,$$

$$(4) \quad y(P) = PQ + x(P_f - P) - zk \quad \text{if } P \geq P_e$$

by introducing the indicator variable I . It is clear from these expressions that sufficient condition (c) for adequacy of the standard MV framework is violated in the presence of options. For example, if $Q - x > 0$ and $Q - x - z < 0$, then $\partial y / \partial P < 0$ for all $P < P_e$ and $\partial y / \partial P \geq 0$ for all $P \geq P_e$. Income is not a monotonic transformation of price; it may be a negative or a positive transformation, which violates condition (c).³

The DM is assumed to believe that output price is normally distributed with a known mean μ_P and variance σ_P^2 . Because P is normally distributed and $y(P)$ is a linear function of P , if $P < P_e$ the unconditional mean and variance of income are [from (3)]

$$\mu_{y1} = (Q - x - z)\mu_P + xP_f + z(P_e - k),$$

$$\sigma_{y1}^2 = (Q - x - z)^2 \sigma_P^2.$$

If $P \geq P_e$, $y(P)$ is normally distributed with unconditional mean and variance μ_{y2} and σ_{y2}^2 [from (4)],

$$\mu_{y2} = (Q - x)\mu_P + xP_f - zk,$$

$$\sigma_{y2}^2 = (Q - x)^2 \sigma_P^2.$$

Define $y(P_e)$ as the value of $y(P)$ from (3) or from (4) with $P = P_e$. The income density function corresponding to equation (2) can be shown to be⁴

$$(5) \quad w(y) = I_1 \exp [-(y - \mu_{y1})^2 / 2\sigma_{y1}^2] / \sigma_{y1} \sqrt{2\pi} + I_2 \exp [-(y - \mu_{y2})^2 / 2\sigma_{y2}^2] / \sigma_{y2} \sqrt{2\pi},$$

where

$$I_1 = 1 \text{ for } y < y(P_e) \quad \text{if } (Q - x - z) > 0, \\ \text{or for} \\ y \geq y(P_e) \quad \text{if } (Q - x - z) \leq 0, \\ = 0 \text{ otherwise;}$$

² Commodity options currently traded on the organized exchanges are options on commodity futures contracts. The income function in equation (2) treats the options as options on the spot commodity. The assumption of a certain basis level results in equivalence between the two types of options. That is, if the basis is known with certainty, then an option on a futures contract is equivalent to an option on the underlying spot commodity.

³ Sufficient condition (c) is violated when options are included in a portfolio. However, as Meyer points out, it may be difficult to reject the null hypothesis that sufficient condition (c) holds when using empirical data.

⁴ The derivation is presented by Hanson.

$$I_2 = 1 \text{ for } y \geq y(P_e) \quad \text{if } (Q - x) \geq 0, \\ \text{or for} \\ y < y(P_e) \quad \text{if } (Q - x) < 0, \\ = 0 \text{ otherwise.}$$

The first term of (5) corresponds to equation (3); the second, to equation (4). This distribution is the sum of two truncated normal distributions, for which the mean, variance, and bounds of integration depend on μ_p , σ_p^2 , P_f , P_e , k , Q and the DM's choices of x and z . The distribution represented by equation (5) can take on a variety of forms dependent on the DM's positions in the futures and options markets. It is clear from (5) that sufficient condition (b) is violated because the income distribution is not normal when the DM takes a position in the options market. Thus, sufficient conditions (b) and (c) for the standard MV model to produce results that are consistent with expected utility maximization are violated when options contracts are available, and an alternative framework is needed.

Expected Utility Maximization

If the DM is assumed to maximize expected utility and faces the income function and conditions indicated by equation (2), then the DM chooses futures and options positions to satisfy equation (6):

$$(6) \quad \text{MAX}_{x,z} \bar{U} = \int_{-\infty}^{\infty} U(y)w(y)dy,$$

where \bar{U} is the DM's expected utility given the choice of x and z ; $U(y)$ is the DM's utility function, assumed to exhibit CARA; and $w(y)$ is the income distribution specified by equation (5). Solving equation (6) yields the optimal futures and options positions, x^* and z^* . Because equation (6) has no explicit algebraic solutions when the DM uses options, solution techniques involving numerical integration and optimization are required.

Optimal Hedging in a Mean-Variance Framework

Expression (5) shows that the assumptions underlying the linear MV model are violated when options are held by the DM. The linear MV model may, however, produce acceptable quantitative results or useful insights into the impacts

of options on optimal market positions. To investigate this possibility, this section extends the linear MV model to include options.

Three different MV models are used to study the optimal market positions under various situations. Model 1 presents the standard result for a DM who uses only the cash and futures markets and corresponds to the use of a security which does not truncate the income distribution.⁵ Model 2 includes the cash and options market and corresponds to the use of a security which truncates the income distribution. Model 3 includes the cash, futures and options markets and corresponds to the use of a security which truncates the income distribution and a security that does not truncate the income distribution. Model 1 contains no truncation problem but models 2 and 3 do.⁶ Each model is solved using (1), and implications of the solutions are discussed. As will be seen later, despite the violation of the sufficient conditions for MV consistency with EU, the model does provide useful insights into the numerically generated EU results.

These models use the following additional definitions: T is localized value of the option at the end of the period = $\max [0, P_e - P]$; μ_T is expected value of the option; σ_T^2 is variance of the value of the option; $\sigma_{p,T}$ is covariance between the value of the option and the spot price.

Model 1: Cash and Futures with Certain Output (CF)

Model 1 examines the optimal hedge in a futures market for a DM having a fixed end-of-period output level. The end-of-period income function is

$$(7) \quad y = PQ + x(P_f - P).$$

Solving (1) subject to (7) yields (8) as the optimal futures position

$$(8) \quad \hat{x} = Q + \frac{(P_f - \mu_p)}{A\sigma_p^2}.$$

It is useful to discuss the results in each MV model in terms of the hedging and speculative components. The first term in the solution (8) is the hedging component; it is simply the futures position that minimizes the variance of end-of-period income. The second term is the spec-

⁵ These results have been reported in previous work: for example, see Anderson and Danthine.

⁶ Models 2 and 3 are derived in Hanson.

ulative component; it is the futures position that results if the DM does not have any cash position and obtains his entire end-of-period income from the futures market. If the futures market is considered unbiased (i.e., $P_f = \mu_p$), then the speculative component disappears, and the optimal futures position is the traditional hedge where the futures position is equal and opposite the cash position. However, if the DM believes the futures market is biased, then the DM believes that, on average, profits can be made by speculating in the futures market. The futures position is altered by the amount of believed bias adjusted by the level of risk aversion and price variability. Increases in risk aversion lead to smaller speculative positions. If a DM is infinitely risk averse, the speculative component is zero, and the futures position equals the traditional hedge.

Model 2: Cash and Options with Certain Output (CO)

Model 2 examines the options market position. End-of-period income consists of revenue from the cash and options markets less the cost of purchasing options. The income function can be written as

$$(9) \quad y = PQ + zT - zk,$$

where both P and T are random variables. Choosing z to maximize (1) subject to (9) yields

$$(10) \quad \hat{z} = -\frac{Q\sigma_{p,T}}{\sigma_T^2} + \frac{(\mu_T - k)}{A\sigma_T^2}.$$

The first and second terms represent hedging and speculative components analogous to model 1. It can be shown that the absolute value of the ratio $\sigma_{p,T}/\sigma_T^2$ will always be greater than one. This implies that in the absence of any speculative position, the DM will always overhedge in the options market. The result occurs because the overhedge position in the options market minimizes the variance of end-of-period income. The speculative component will modify the optimal market position in a manner analogous to model 1.

Model 3: Cash, Futures, and Options with Certain Output (CFO)

Model 3 considers a DM who is able to use futures and options contracts. Income now in-

cludes revenue from the cash, futures, and options markets, and the cost of purchasing options, and can be expressed as

$$(11) \quad y = PQ + x(P_f - P) + zT - zk.$$

Maximizing (1) subject to (11) yields the optimal market positions in the futures and options markets:

$$(12) \quad \hat{x} = Q - \frac{(P_f - \mu_p)\sigma_T^2 + (\mu_T - k)\sigma_{p,T}}{A(\sigma_{p,T}^2 - \sigma_p^2\sigma_T^2)},$$

$$\hat{z} = -\frac{(P_f - \mu_p)\sigma_{p,T}^2 + (\mu_T - k)\sigma_p^2}{A(\sigma_{p,T}^2 - \sigma_T^2\sigma_p^2)}.$$

The optimal futures market position consists of a hedging and speculative component, and the optimal options market position simply consists of a speculative component. In the absence of any speculative positions, the optimal market positions become the traditional hedge in the futures market and no position in the options market. The DM is acting to minimize variance of the end-of-period income, which can be completely eliminated by taking the traditional hedge in the futures market. If either the futures market or options market is perceived to be biased, then the speculative components adjust the futures and options positions accordingly.

Opportunity Cost of MV Solution

The opportunity cost to the DM of using the MV market position instead of the EU position is the amount of additional income that must be given to the DM to provide the same level of expected utility as would be provided by using the EU market position. This money value, which we denote by V , is obtained by solving equation (13) for V .

$$(13) \quad EU(y(\hat{x}, \hat{z}) + V) = EU(y(x^*, z^*)),$$

where \hat{x} and \hat{z} are the MV market positions and x^* and z^* are the EU market positions.

Three dimensions of the MV and EU solutions are compared: (a) value of V , (b) correlation between x^* and \hat{x} , and (c) correlation between z^* and \hat{z} . For the DM studied here, the value of V is the only relevant measure of goodness of fit of MV to EU solutions. Large variations in options and futures positions are economically important only if they affect $E(U)$. It is possible that wide differences in the values of

the decision variables between the MV and EU frameworks do not make much difference in the level of the DM's expected utility.

Experimental Design

Equations (5) and (6) were used to derive optimal futures and options positions of a representative midwestern soybean producer who exhibits constant absolute risk aversion (CARA) subject to the preceding assumptions. These EU solutions were compared with corresponding MV solutions from (1). We could solve (6) for historical values of the market factors. But because commodity options have existed for a short time, this would provide a small sample of results. Consequently, we elected to solve (6) for a simulated set of market factors. To efficiently extract information from a given set of market factors, the factors were organized in a modified factorial design.

Values of the market factors were selected to represent conditions in midwestern soybean markets. The exogenous factors of the model were specified in terms of the levels of four "market factors," amount of output, and the level of risk aversion exhibited by the producer. The four market factors are: σ_p^2 , the variance of the soybean cash price; $B_f = P_f - \mu_p$, the perceived expected bias in the futures market; $B_o = E[P_e - P|P < P_e] - K$, the perceived expected bias in the options market; and $E_m = P_e - \mu_p$, the difference between the exercise price and the expected soybean cash price. Setting the exogenous factors of the model— σ_p^2 , μ_p , P_e , P_f , and k —in relation to these market factors ensured that various combinations of variable values in the design did not result in unrealistic specifications of market behavior.

The variance of average deflated November spot prices received by Iowa soybean producers during 1976–86 was \$1.08. We assigned values of σ_p^2 ranging from \$0.50 to \$1.50. From studies of pricing in futures markets (Gray; Just and Raussier; Conorella and Pollard; Telser; Cootner 1977a, b) and in options markets (Tucker, Whaley, Jordan et al., Luft and Fielitz), we conclude that, if any bias does exist in either futures or options markets, it is small and short-lived. The extreme values of each bias were set at $\pm\$0.04$. The extreme values of the difference between the exercise price and the expected spot price were set at $\pm\$0.10$.

The assigned values of the market factors and of Q and A were used in a central composite

experimental design (Khuri and Cornell). A factorial design is one in which all values of a given factor are combined with all levels of every other factor in the data set. A central composite design consists of a full factorial design for two levels of variable values called the upper and lower factorial levels and additional selected combinations of variable values around the two-level factorial design called the center and axial points. The intent of the central composite design is to capture much of the information of a full factorial design that uses more than two levels of each variable without the data requirements of a full three-level, or higher, factorial design. Table 1 presents the levels of the factors used in the design. The levels in this table were used in 77 different combinations to form a central composite design of $n = 77$ points.

To solve (6) we then applied MINTDF numerical integration (Kaylen and Preckel) and GQOPT/PC (Quandt and Goldfeld) numerical optimization to each combination of variable values to find the optimal futures and options positions, x^* and z^* .

Comparison of MV and EU Results

The MV results from (1) and the EU results from (6) were found for each combination of variables in the central composite design subject to the income functions in (7), (9), and (11). Thus, the MV solution could be directly compared to the EU solution for each model. The general behavioral characteristics of the EU results correspond to the economic behavior described by the MV model in the previous section. As the MV suggested, the EU maximizer hedges to reduce income variability and speculates according to any believed bias in the commodity markets adjusted by risk aversion and income variability.

The MV and EU results were identical for the CF model because the necessary and sufficient condition is met, which causes the linear MV model to produce the EU maximizing solution. As expected, the CO and CFO results for the MV model were different from those for the EU model because of the violation of sufficient conditions for equivalence between the MV and EU models. However, visual examination of the results of the MV and EU models suggests that the MV model does approximate the EU results. The twelve tabular pages of MV and EU solutions are not presented here but are available from the authors.

Table 1. Observation Levels in the Central Composite Design

Levels	Factors					
	E_m	σ^2	B_f	B_0	Q	A
	(\$)				(bu.)	
Upper axial	0.10	1.50	0.04	0.04	25,000	0.00045
Upper factorial	0.05	1.25	0.02	0.02	20,000	0.00035
Center point	0.00	1.00	0.00	0.00	15,000	0.00025
Lower factorial	-0.05	0.75	-0.02	-0.02	10,000	0.00015
Lower axial	-0.10	0.50	-0.04	-0.04	5,000	0.00005

Cash and Options

Results for the CO model are summarized in tables 2 and 3 and in the regression

$$z^* = 1,602 + 0.874 \hat{z} + e, \\ (313) \quad (0.015)$$

where $r^2 = 0.978$, the standard deviation of $e = 847$ bushels, and the values in parentheses are the standard errors of the estimated coefficients. The standard deviation of the errors amounted to only 17% of the smallest value of Q included in the experiment. An F -test rejected the null hypothesis that the intercept equalled zero and the slope equalled one at the 1% level. All values of \hat{z} and z^* were positive.

Table 2 contains the frequency distribution of

Table 2. Distribution of Relative Absolute Differences Between EU and MV Options Positions

Relative Absolute Difference ^a	Frequency		
	CO Options	Futures	Options
(%)			
0-5	8	47	41
5-10	14	15	11
10-15	22	9	10
15-20	7	4	6
20-25	4	2	5
25-30	4	0	2
30-35	0	0	2
35-40	3	0	0
40-45	2	0	0
45-50	0	0	0
50-55	4	0	0
55-60	4	0	0
60-65	3	0	0
65-70	2	0	0
70 and up	0	0	0

^a $100|z^* - \hat{z}|/5,000$ and $100|x^* - \hat{x}|/5,000$. These are absolute differences expressed as percentages of number of bushels in a futures or options contract.

Table 3. Frequency Distribution of the Opportunity Cost of MV Market Positions

Value V	Frequency	
	CO Model	CFO Model
(\$)		
0-2 ^a	13	41
2-4	13	10
4-6	4	6
6-8	4	8
8-10	1	5
10-12	3	0
12-14	2	4
14-16	1	1
16-18	2	2
18-20	6	0
20-100	12	0
100-200	8	0
200-300	6	0
300-350	2	0
Mean	\$54.66	\$3.80
Median	\$12.94	\$1.50
Mode	\$ 3.50	\$0.50

^a Includes lower limit and excludes upper limit on range.

the relative absolute differences (RADs) between the EU and MV market positions: i.e., the absolute differences divided by 5,000 bushels, the size of a standard futures or options contract. The majority of the observations (57%) have a RAD of less than 15%, while the remaining observations are spread uniformly up to a RAD of 70%. The largest error in the MV solutions was 3,366 bushels, which amounted to 14.6% of the EU solution. Table 3 contains the frequency distribution of the opportunity costs of the MV market positions. The mean is around \$55 and the median is about \$13. The distribution is skewed with a range of nearly zero to \$350. The largest concentration of observations (64%) is in the \$0-\$20 range. Of the seventy-seven values of V , sixty-eight are less than 1¢ per bushel of Q . The largest value of V amounts of \$342, or slightly less than 2¢ per bushel.

Cash, Futures, and Options

Results for the CFO Model are summarized in tables 2 and 3 and in the simple regressions:

$$z^* = 51 + 1.68 \hat{z} + e, \quad (23) \quad (0.03)$$

where $r^2 = 0.978$, and the standard deviation of $e = 201$ bushels. An F -test rejected the hypothesis of a zero intercept and a unit slope at the 1% level.

$$x^* = -113 + 1.005 \hat{x} + e, \quad (153) \quad (0.010)$$

where $r^2 = 0.993$, and the standard deviation of $e = 412$ bushels. An F -test failed to reject the hypothesis of a zero intercept and a unit slope at the 40% level. The standard deviations of the errors in these two equations are less than 10% of the smallest value of Q included in the experiment. Some values of \hat{z} and z^* are positive, and some negative. Every value of z^* is less than 5,000 bushels (one options contract) and only eight exceed 2,500 bushels.

Table 2 shows that all of the RADs in the futures positions are less than 25% and 61% have a RAD of 5% or less. The RADs of the options positions are all less than 35%, and 53% of the observations have a RAD of less than 5%. The largest errors in \hat{x} and \hat{z} are 1,159 bushels and 1,628 bushels, which are 14% and 51% of the correct values x^* and z^* . The futures position has a relatively low percentage error because the traditional hedge is in the futures market in MV and EU models. The entire error in the options position is related to the speculative component. The resulting speculative error in the futures position for this particular observation was 61%.

Table 3 shows the opportunity cost to the DM of using the MV market positions instead of the EU positions has a mean of \$3.80 and a median of \$1.50. The distribution is skewed with all observations falling in the range \$0–\$18. The largest concentration of observations (53%) is in the \$0–\$2 range. The largest value of V is \$16.38, less than 0.2¢ per bushel of output.

Values of V and \hat{z} are positively correlated, as are values of V and A . As the size of the options position in the expected utility solution increases, the error in the standard MV solution goes up, and the value of the EU decision over the MV solution also rises. As the options position increases, the resulting income distribution deviates further from the normal distribution assumed by the MV model. Consequently,

the MV model results contain more error as the income distribution becomes "less normal." As the level of risk aversion rises, any deviation from the optimal solution is translated into a higher opportunity cost being attached to the MV solutions.

Comparative Statics: MV versus EU

The previous results show that the MV model can be used to produce results that approximate the solutions of the EU model. The analytical results of the MV model are also useful in providing insights into the numerical results generated by the EU model. For example, in EU solutions to the CO model the optimal options position is always greater than the output level. Using the analytical result from (10), it is easy to show that the variance of end-of-period income is minimized by taking an options position that is greater than the output level.

One test of the MV model's analytical usefulness is its ability to produce comparative static results consistent with those produced by the EU model. The comparative static results can be generated for the MV model directly by using equations (10) and (12). In some cases, the MV model comparative static results cannot be signed in general; e.g., $\partial x^*/\partial \sigma^2$ and $\partial z^*/\partial \sigma^2$ in the CFO strategy. However, the comparative static results for both the MV and EU models can be generated numerically over local regions. Table 4 presents the local comparative results over the region considered in this study. The comparative static results of the MV model are not always consistent with the EU model. In the CO strategy, there is some inconsistency between signs of $\partial x^*/\partial \sigma^2$ and $\partial x^*/\partial A$ across the various combinations of exogenous variables. In the CFO strategy there is some inconsistency in the two models between $\partial x^*/\partial E_m$ and $\partial z^*/\partial E_m$. The remaining comparative static results are consistent between the two models over the region considered.

The standard linear MV model used in this analysis is a special MV model. In general, the form of the MV function will not be known. It may be that the EU choices would be in the MV efficient sets generated by some alternative non-linear MV preference functions. In an alternative specification of the MV preference function, the differences between the two sets of results may be even smaller. Relaxing the assumptions such as CARA, normally distributed output price, no transactions costs, no basis un-

Table 4. Numerical Comparative Static Results for the MV and EU Models

Δ Factor ^a	CO			CFO			
	Δ Options			Δ Futures		Δ Options	
	MV	EU		MV	EU	MV	EU
			Special Conditions				
$\Delta E_m > 0$	< 0	< 0	$B_f = B_0 > 0$	> 0	≈ 0	< 0	< 0
			$B_f = B_0 < 0$	< 0	≈ 0	> 0	> 0
			$B_f > 0, B_0 < 0$	< 0	> 0	> 0	< 0
			$B_f < 0, B_0 > 0$	> 0	< 0	< 0	> 0
$\Delta \sigma^2 > 0$	≈ 0	≈ 0	$B_f = B_0 > 0$	> 0	> 0	< 0	< 0
			$B_f = B_0 < 0$	< 0	< 0	> 0	> 0
			$B_f > 0, B_0 < 0$	< 0	< 0	> 0	> 0
			$B_f < 0, B_0 > 0$	> 0	> 0	< 0	< 0
$\Delta B_f > 0$	$= 0$	$= 0$		> 0	> 0	< 0	< 0
$\Delta B_0 > 0$	> 0	> 0		< 0	< 0	> 0	> 0
$\Delta Q > 0$	> 1	> 1		$= 1$	$= 1$	$= 0$	$= 0$
	Special Conditions						
$\Delta A > 0$	$B_0 > 0$	< 0	$B_f = B_0 > 0$	> 0	> 0	< 0	< 0
	$B_0 < 0$	> 0	$B_f = B_0 < 0$	< 0	< 0	> 0	> 0
			$B_f > 0, B_0 < 0$	< 0	< 0	> 0	> 0
			$B_f < 0, B_0 > 0$	> 0	> 0	< 0	< 0

^a The symbol Δ is used to represent a change in the indicated variable.

certainty, and no margin calls may also change some of the results. The impact of these assumptions is left for further study.

Conclusions and Implications

This study considered the approximating ability of the standard linear MV model for a DM with constant absolute risk aversion when the income distribution is truncated by the use of commodity options. The MV model was shown to provide analytical results useful in describing the general behavioral characteristics of the EU model. However, the comparative static results of the MV model were not always consistent with the comparative static results of the EU model. Thus, some care must be taken in interpreting the analytical results of the MV model when the sufficient conditions for expected utility equivalence are violated.

The MV model also produced empirical solutions that were close approximations of the EU model results. The opportunity costs of using the mean-variance solutions instead of the EU solutions were small. In the situations studied here, the payoff for performing the additional computations (the numerical integration and numerical optimization) instead of using the simpler MV solution is small. The MV model is robust to the particular violation of the assump-

tions that was studied here. These results will also be of interest to those concerned with modeling other decision-making problems under risk which are characterized by truncated probability distributions.

[Received August 1989; final revision received September 1990.]

References

- Anderson, Ronald W., and Jean-Pierre Danthine. "Hedger Diversity in Futures Markets." *Econ. J.* 93(1983):370-89.
- Baron, D. "On the Utility Theoretic Foundations of the Mean-Variance Analysis." *J. Finan.* 32(1977):1683-97.
- Conorella, G., and S. K. Pollard. "Efficiency of Commodity Futures: A Vector Autoregressive Analysis." *J. Futures Mkts.* 5(1985):57-76.
- Cootner, Paul H. "Rejoinder." *Selected Writings on Futures Markets, Volume II*, ed. A. E. Peck, pp. 65-69. Chicago: Chicago Board of Trade, 1977b.
- . "Returns to Speculators: Telser vs. Keynes." *Selected Writings on Futures Markets, Volume II*, ed. A. E. Peck, pp. 41-51. Chicago: Chicago Board of Trade, 1977a.
- Gray, R. W. "The Characteristic Bias in Some Thin Futures Markets." *Selected Writings on Futures Markets, Volume II*, ed. A. E. Peck, pp. 113-30. Chicago: Chicago Board of Trade, 1977.
- Hanson, Steven D. "Price Level Risk Management in the

- Presence of Commodity Options: Income Distribution, Optimal Market Positions, and Institutional Value." Ph.D. thesis, Iowa State University, 1988.
- Jordan, J. V., W. E. Seale, N. C. McCabe, and D. E. Kenyon. "Transactions Data Tests of the Black Model for Soybean Futures Options." *J. Future Mkts* 7(1987):535-54.
- Just, R. E., and G. C. Rausser. "Commodity Price Forecasting with Large-Scale Econometric Models and the Futures Market." *Amer. J. Agr. Econ.* 63(1981):197-208.
- Kahl, K. H. "Determination of the Recommended Hedging Ratio." *Amer. J. Agr. Econ.* 65(1983):603-5.
- Kaylen, Michael, and Paul Preckel. "MINTDF User's Guide." Staff Pap. No. AEWP 1986-22, Purdue University, 1986.
- Khuri, Andre I., and John A. Cornell. *Response Surfaces: Designs and Analyses*. New York: Marcel Dekker, 1987.
- King, R. P., and L. J. Robison. "An Interval Approach to the Measurement of Decision Maker Preferences." *Amer. J. Agr. Econ.* 63(1981):510-20.
- Levy, H., and H. M. Markowitz. "Approximating Expected Utility by a Function of Mean and Variance." *Amer. Econ. Rev.* 69(1979):308-17.
- Luft, Carl F., and B. C. Fielitz. "An Empirical Test of the Commodity Option Pricing Model Using Ginnie Mae Call Options." *J. Finan.* 41(1986):137-51.
- Markowitz, H. "Portfolio Selection." *J. Finan.* 7(1952):77-91.
- Meyer, Jack. "Two Moment Decision Models and Expected Utility Maximization." *Amer. Econ. Rev.* 77(1987):421-30.
- Meyer, Jack, and Robert H. Rache. "Kolmogorov-Smirnov Tests for Distribution Function Similarity with Applications to Portfolios of Common Stock." New York: National Bureau of Economic Research Tech. Work. Pap. No. 76, 1989.
- Meyer, Jack, and Lindon J. Robison. "Hedging Under Output Price Randomness." *Amer. J. Agr. Econ.* 70(1988):268-72.
- Musser, Wesley N., Harry H. Mapp, Jr., and Peter J. Barry. "Applications I: Risk Programming." *Risk Management in Agriculture*, ed. Peter J. Barry, chap. 10. Ames: Iowa State University Press, 1984.
- Nelson, Carl H. "Elliptical Symmetry and Mean-Variance Portfolio Choice." Proceedings of the S-232 Regional Research Committee on Quantifying Long-Run Agricultural Risks and Evaluating Farmer Responses to Risk. Dep. Food and Resource Econ., University of Florida, April 1990.
- Newbery, David M., and Joseph E. Stiglitz. *The Theory of Commodity Price Stabilization*. New York: Oxford University Press, 1981.
- Porter, R. B. "An Empirical Comparison of Stochastic Dominance and Mean-Variance Portfolio Choice Criteria." *J. Finan. and Quant. Anal.* 8(1973):587-608.
- Quandt, Richard E., and Stephen M. Goldfeld. "GQOPT/PC Version 3.61." Econ. Dep., Princeton University, 1987.
- Robison, Lindon J., and Peter J. Barry. *The Competitive Firm's Response to Risk*. New York: Macmillan Co., 1987.
- Robison, Lindon J., and John R. Brake. "Application of Portfolio Theory to Farmer and Lender Behavior." *Amer. J. Agr. Econ.* 61(1979):158-64.
- Samuelson, P.A. "The Fundamental Approximation Theorem of Portfolio Analysis in Terms of Means, Variances and Higher Moments." *Rev. Econ. Stud.* 37(1970):537-42.
- Telser, Lester G. "Reply." *Selected Writings on Futures Markets, Volume II*, ed. A. E. Peck, pp. 52-64. Chicago: Chicago Board of Trade, 1977.
- Tew, Bernard V., and Donald W. Reid. "More Evidence on Expected Value-Variance Analysis Versus Direct Utility Maximization." *J. Finan. Res.* 10(1987):249-57.
- Tobin, J. "Liquidity Preference as Behavior Toward Risk." *Rev. Econ. Stud.* 37(1958):65-86.
- Tsiang, S. C. "The Rationale of the Mean-Standard Deviation Analysis, Skewness Preference, and the Demand for Money." *Amer. Econ. Rev.* 62(1972):354-71.
- Tucker, A. L. "Empirical Tests of the Efficiency of the Currency Options Market." *J. Finan. Res.* 8(1985):275-85.
- Whaley, R. "Valuation of American Futures Options: Theory and Empirical Test." *J. Finan.* 41(1986):127-50.
- Wolf, Avner. "Optimal Hedging with Futures Options." *J. Econ. and Bus.* 39(1987):141-58.

Evaluating Robust Regression Techniques for Detrending Crop Yield Data with Nonnormal Errors

Scott M. Swinton and Robert P. King

Although ordinary least squares is not efficient when errors are not distributed normally, it generates better crop yield trend coefficient estimates than six alternative robust regression methods. This is because of the econometric properties of an uninterrupted series independent variable as well as the level of skewness typical of corn yields. The evaluation covers actual farm-level corn yield series as well as a set of "contaminated" data series and one thousand sets of Monte Carlo yield series. Where an influential end-of-series outlier is suspected, the DFBETAS regression diagnostic statistic is recommended.

Key words: detrending, regression diagnostics, robust estimation, yield distributions.

When detrending farm-level crop yield time-series data, a series-end observation which lies far from the regression trend can exert undue influence on the trend coefficient. Until recently, the standard solution was to omit such an outlier from the data series. Now, however, a growing number of robust regression techniques make it possible to weight observations automatically so that outliers have reduced influence or none at all. This paper demonstrates that ordinary least squares (OLS) regression remains the preferred means of detrending time-series corn yield data when compared with a set of robust regression (RR) techniques. Conclusions are drawn from comparisons using (a) three sets of farm-level corn yields, (b) varying lengths of farm yield time series with an artificial series-end outlier, and (c) one thousand sets of randomly generated artificial corn yield data in time series of ten to twenty-five periods.

Crop yield time-series data are detrended in order to "remove" technology bias from estimates of the underlying probability distribution. Among other purposes, crop yield distributions are used in evaluating alternative crop plans and in analyses of crop insurance decisions. For

practical purposes, such distributions should be estimated by a method that (a) is simple, unbiased, and reasonably efficient, (b) requires only yield and time data, and (c) can accommodate short data series.

Detrending is generally accomplished by regressing observed values of the yield random variable on some measure of time using least squares methods. This approach yields maximum likelihood estimates under the assumption that the regression residuals are normally distributed. While this assumption may be suitable for some purposes, empirical tests have found that it often fails to hold for agronomic and farm-level crop yield data (Day, Nelson and Preckel). The negative skewness typical of corn yield probability density functions (pdf) implies that low extreme values are more likely to occur than if the pdf were normal. For example, a low, end-of-series outlier occurred in the Midwest in 1988 as a result of drought.

By down-weighting outlier observations, RR techniques may generate more efficient trend coefficient estimates when the data are not normally distributed. Such methods have been discussed extensively in the statistics literature during the past twenty years and are becoming increasingly available to applied economists in econometric software packages. However, it will be shown that they do not necessarily offer coefficient estimates that differ significantly from OLS estimates, even when errors are not normally distributed. This paper will (a) illustrate

Scott M. Swinton is a research assistant and Robert P. King is a professor, Department of Agricultural and Applied Economics, University of Minnesota.

Contribution No. 17,215 of the Minnesota Agricultural Experiment Station based on research supported by the station and by a grant from the Federal Crop Insurance Corporation.

cases in which several robust regression methods on an equation having nonnormal errors fail to give coefficient estimates significantly different from OLS and (b) demonstrate how regression diagnostics can identify conditions under which RR is a useful alternative to OLS for detrending time-series data.

Nonnormal Errors and Robust Regression

In the standard linear model, with n observations and k regressors,

$$Y = X\beta + u,$$

where Y is an $(n \times 1)$ dependent variable vector, X is an $(n \times k)$ matrix of k independent variables, β is a $(k \times 1)$ coefficient vector, and u is an $(n \times 1)$ disturbance vector, OLS will yield maximum likelihood estimates of model parameters when $u \sim N(0, \sigma^2)$. When u is not distributed normally, the OLS coefficient estimator, b , is still the best linear unbiased estimator, and the OLS variance estimator, s^2 , is still unbiased and consistent. However, neither b nor s^2 is efficient or asymptotically efficient, since the maximum likelihood estimator is nonlinear (Judge et al., p. 888).

Because OLS minimizes the sum of squared errors, large residuals or "outliers" can exert considerable influence on parameter estimates. Robust regression methods give less weight than OLS to observations whose dependent variable deviates far from its expected value. They can be viewed as automated means of reducing outlier influence and are classified by their approach to controlling influential outliers. M -estimators employ maximum-likelihood techniques for finding regression coefficients that minimize some function of the regression residuals, typically one that gives less weight than OLS to residuals large in absolute value. Linear combinations of order statistics, or L -estimators, calculate regression coefficients for quantiles of the residuals resulting from a regression model and then combine them with specified weights.

Both classes of RR methods were used to detrend three corn yield time series after discovering that the OLS residuals were negatively skewed in two of the series. The structural model for each farm took the simple, linear form:

$$CORN\text{YLD} = \beta_0 + \beta_1 \cdot \text{YEAR} + u,$$

where $CORN\text{YLD}$ denotes the corn grain yield in bushels per acre, and YEAR is the corresponding year. The three series came from farms

in Jackson County, Minnesota, having 22 to 43 annual yield observations ending between 1985 and 1987. Due to the negative skewness of the OLS residuals from farms A and B, it was suspected that outliers might be distorting the OLS coefficient estimates. Six RR methods discussed in Judge et al. were used to reestimate the trend parameter: the multivariate t , which is an M -estimator, and five L -estimators, the least absolute error (LAE), trimmed mean ($TRIM$), five quantity-weighted regression quantile ($FIVEQUAN$), the Gastwirth weighted regression quantile ($GASTWIRTH$), and Tukey tri-mean weighted regression quantile ($TUKEY$) procedures, available in SHAZAM (White, Shirley, and Horsman).

As summarized in table 1, all estimates of the coefficient on YEAR lie within one OLS standard error of the OLS estimate. For all three farms, coefficient estimates with RR are not significantly different from OLS, despite (1) a disproportionately large number of negative residuals for farms A and B, (2) negative summed residuals on all three farms for all RR methods used, and (3) a relatively large OLS standard error for farm C.

Regression Diagnostics

Three regression diagnostic statistics are particularly helpful in gauging the influence of an ob-

Table 1. Coefficient Estimates Compared for YEAR Variable in the Linear Model: OLS versus Six Robust Regression Methods

Regression Method	Farm A	Farm B	Farm C
Parameters			
Degrees of freedom (d.f.)	40	41	20
OLS standard error of coefficient	0.214	0.242	0.612
Coefficient estimates			
OLS	2.034	2.062	1.503
LAE	2.166	2.203	1.512
TRIM = .05	2.039	2.209	*
TRIM = .10	2.049	2.184	1.562
TRIM = .20	2.129	2.086	1.447
FIVEQUAN	2.073	2.033	1.836
GASTWIRTH	2.068	2.083	1.322
TUKEY	2.108	2.091	1.839
MULTIT = 1	2.133	2.149	1.463
MULTIT = 3	2.097	2.127	1.480
MULTIT = d.f.	2.042	2.075	1.498

* This L -estimate could not be computed due to small sample size relative to the size of the desired trim quantile.

servation upon a coefficient estimate: (a) h_i , the measure of potential leverage by individual observations on the independent variables to influence the regression (b) e_i^* , the "studentized" residual, a measure of what constitutes an influential outlier, and (c) $DFBETAS$, a newer diagnostic that measures the impact on a given coefficient of a given observation (Belsley, Kuh, and Welsch).

The potential influence, or leverage, of an observation, h_i (Belsley, Kuh, and Welsch; Judge et al.; Weisberg), is defined as follows:

$$h_i = x_i'(X'X)^{-1}x_i,$$

where x_i is an observation on the independent variable(s) and X is the matrix of all observations on the independent variable(s). Key characteristics of h_i are (a) it always lies between zero and one, (b) the h_i 's sum to k , the number of regressors, and (c) h_i is a function of the independent variables only. As observations on the independent variables, x_i , get farther from the sample mean, they become potentially more influential, and h_i grows larger. Since the average value of h_i is k/n , a conservative rule of thumb for identifying observations with high leverage is $h_i \geq 3k/n$ (Judge et al., p. 893).

A widely employed, reliable measure of whether an extreme residual is indeed a statistical outlier is the "studentized" residual (Belsley, Kuh, and Welsch; Weisberg), e_i^* ,

$$e_i^* = \frac{e_i}{s(i)^*(1 - h_i)^{.5}},$$

where e_i is the residual corresponding to the i th observation and $s(i)$ denotes the sample standard error of the estimate calculated omitting the i th observation. As is clear from its composition, e_i^* will be large if one or more of the following conditions holds: (a) e_i is large, (b) h_i is large, or (c) $s(i)$ is small. The studentized residual follows the central t -distribution with $n-k$ degrees of freedom. However, since it describes a residual which represents one of many "draws" from the distribution, the appropriate test statistic is the Bonferroni t -value, which tests the hypothesis that the residual occurs with a/n probability, where a is the probability of mistakenly rejecting the hypothesis that e_i is an outlier and n is the number of observations in the sample (Weisberg, p. 116).

Belsley, Kuh, and Welsch have developed a statistic called $DFBETAS$ to measure the influence that an observation is likely to have on the regression coefficient. It measures the difference between estimates for the j th coefficient

with (b_j) and without ($b_j(i)$) the i th observation, as standardized by the coefficient standard error:

$$DFBETAS_{ij} = \frac{b_j - b_j(i)}{s(i)[(X'X)_{jj}^{-1}]^{.5}},$$

where $(X'X)_{jj}$ denotes the j th diagonal element of the $(X'X)$ projection matrix. Belsley, Kuh, and Welsch demonstrate that $DFBETAS$ decreases with sample size at a rate proportionate to $n^{-.5}$. Hence, they recommend a "size-adjusted cutoff" of $|DFBETAS| > 2/(n^{.5})$ (p. 28).

Diagnosis of Residuals from Regressions to Detrend Time Series

In a regression to detrend an uninterrupted time series, the time measure increases by a single unit from one observation to the next. This property has distinct consequences for regression diagnostics.

First, the measure of potential leverage, h_i , cannot become very large because there are no x_i values far from the mass of x_i 's. The h_i values are greatest at the beginning and end of the series, following a symmetric U-shaped pattern. Since the h_i 's sum to the number of regressors, in a simple regression they sum to 2. As the number of sequential observations increases, the potential leverage of any one decreases. For uninterrupted time series, h_i values will be the same for any two samples of the same size, and it can be shown that none will exceed the conservative Judge et al. cutoff value of $3k/n$ (Swinton and King, p. 9).

Barring a high h_i value, it is clear that an externally studentized residual can become "large" only if (a) $s(i)$ is small, or (b) e_i is large. However, given the routine variability of much agricultural data, $s(i)$ and e_i do not often reach orders of magnitude great enough for e_i^* to become significant. Furthermore, while any dependent variable outlier may influence the intercept coefficient, only those corresponding to the beginning or ending of the time series will have a marked effect on the slope coefficient.

The most reliable regression diagnostic for revealing the influence of an observation on coefficient estimates is $DFBETAS$. This becomes especially evident in a test of regression diagnostics and robust methods on a "contaminated" data set. A set of eight corn yield data series of five to forty years, all ending in 1984 from the same southwest Minnesota farm, were modified by replacing the 1984 value with an artificial yield observation of 72.7 bushels per acre, three stan-

dard errors below the expected yield of 122.5 bushels per acre. The test was conceived to model the effect that might be expected from low yields caused by severe drought. Table 2 compares regression diagnostics from the 1984 observation and presents estimates of the coefficient on *YEAR*. Of the three regression diagnostic measures, only *DFBETAS* signaled a potential problem with the 1984 observation. The h_i values all remained predictably below the size-adjusted cutoff value. The studentized residuals were small for small samples because the large outlier significantly influenced the regression, reducing e_i and increasing $s(i)$. None of the studentized residuals exceeded its Bonferroni critical t -value.

Although robust regression has been recommended as a pre-diagnostic technique (Weisberg, p. 253), table 2 demonstrates that robust methods are not foolproof. Only the 20% trimmed mean regression procedure generated coefficients consistently within two coefficient standard errors of the OLS estimate on the uncontaminated sample [$1.88 \pm 2(0.23)$]. However, in *SHAZAM*, trimmed mean least squares (TLS) estimates cannot be computed for sample sizes too small to allow at least three observations in the quantiles to be cut from each tail. The least absolute errors (LAE) estimator performed very poorly, and the estimator following the multivariate t -distribution with one degree of freedom

did no better than the OLS estimator. Of the three RR estimators, TLS performed best, generating coefficient estimates closer to the uncontaminated OLS estimates than OLS for $n < 25$. However, TLS could not be computed for the five observation case. By contrast, since *DFBETAS* revealed that 1984 was a potentially influential observation in every case, its use would have permitted obtaining better OLS estimates by deleting the 1984 observation.

OLS versus Trimmed Least Squares: A Monte Carlo Experiment

In the "contaminated" sample examined above, TLS outperformed both OLS and the other RR methods tested in small samples. A Monte Carlo experiment was conducted to validate these results on the kind of error distributions encountered in detrending farm-level corn yields. Based on the finding of Nelson and Preckel that Iowa farm-level corn yields follow a negatively skewed beta distribution, four sets of time-series data were generated for yields with a trend of 2 bushels per year and a random error drawn from three different beta distributions and one normal distribution, all with mean 0. Since the beta distribution is defined on the interval [0, 1], the generated beta random variates were centered

Table 2. Comparison of Regression Diagnostics and Robust Regression Methods on Corn Grain Yield Detrending Regressions for Farm A with a "Contaminated" 1984 Observation

Number of Observations	Year Series	Studentized Residual for 1984 ^a	<i>DFBETAS</i>	Coefficient on <i>YEAR</i>			
				OLS	LAE	Mt = 1 ^b	TRIM = .2
5	1980-84	-1.02	-1.02 ⁺	-14.32 (5.24) ^d		-17.16	
10	1975-84	-2.30	-1.41 ⁺	1.13 (2.96)	4.61	2.46	1.87*
15	1970-84	-2.10	-1.01 ⁺	1.05 (1.44)	2.59	1.41	1.55*
20	1965-84	-2.52	-1.03 ⁺	2.22 (0.89)	2.67	2.56	1.86*
25	1960-84	-2.53	-0.92 ⁺	1.93* (0.58)	2.67	2.26*	2.00*
30	1955-84	-2.68	-0.88 ⁺	1.91* (0.41)	2.51	2.23*	2.07*
35	1950-84	-2.61	-0.79 ⁺	1.84* (0.32)	2.51	2.13*	2.22*
40	1945-84	-2.63	-0.74 ⁺	1.73* (0.25)	1.96*	1.92*	1.82*

Note: The "contaminated" 1984 observation was three standard errors of estimate below the expected value for that year.

^a Bonferroni critical values exceeded 3.48 in all cases.

^b Denotes multivariate t -distribution with 1 degree of freedom.

^c Plus indicates *DFBETAS* value exceeds size adjusted cutoff of $2/(n^5)$.

^d Standard error in parentheses.

^e Asterisk indicates coefficient estimate lies within two coefficient standard errors of uncontaminated OLS estimate for 1945-84, $1.88 \pm 2(0.23)$.

Table 3. Parameters of the Distributions from Which Random Errors Were Generated

Distribution	Minimum	Maximum	Mean	Standard Deviation
Beta (3, 3)	-50	50	0	18.90
Beta (5, 3)	-62.5	37.5	0	16.14
Beta (15, 3)	-83.3	16.7	0	8.55
Normal (0, 625)	-infinity	infinity	0	25.00

on 0 by subtracting the mean and then scaled over a range of 100. Parameters for the beta distributions were (3, 3), a symmetric distribution; (5, 3), a negatively skewed distribution in the range encountered by Nelson and Preckel; and

(15, 3), a still more negatively skewed distribution. Random errors were also generated from a normal (0, 625) distribution, for which 95% of the distribution falls in the 100-bushel range of the scaled beta (3, 3) distribution. Parameters of distributions from which the random errors were generated are presented in table 3.

Key results from 1,000 regressions using both OLS and trimmed least squares (trim proportion = .2) on randomly generated data from the four distributions reveal that under OLS, trend coefficient estimate ranges are narrower and standard deviations are smaller in all cases, as shown in table 4. In most cases, the sample mean of the trend coefficient estimates is as close or closer to the true value of 2 under OLS as under TLS for data from all four distributions. OLS gen-

Table 4. Summary Statistics on Trend Coefficient Estimates from 1,000 OLS and Trimmed Least Squares (Trim Proportion = .2) Regressions on Data Generated from Four Error Distributions Using Four Sample Sizes

Distribution and Method	Sample Size		Runs	Mean	Standard Deviation	Minimum	Maximum
	Original	After Trim					
Beta (3, 3)							
OLS	10	10	1000	1.8687	2.1061	-5.2377	7.8670
Trim = .2	10	4	641	1.7949	2.9783	-8.5324	15.931
OLS	15	15	1000	2.0004	1.1654	-1.7042	5.4358
Trim = .2	15	7	846	2.0132	1.4683	-2.3152	6.5875
OLS	20	20	1000	1.9787	0.7500	-0.3687	4.1929
Trim = .2	20	10	929	1.9727	0.3390	-1.3164	4.4149
OLS	25	25	1000	2.0103	0.5177	0.6056	3.5099
Trim = .2	25	13	950	2.0040	0.5323	-0.2780	3.8566
Beta (5, 3)							
OLS	10	10	1000	1.9901	1.3894	-3.0522	7.7586
Trim = .2	10	4	635	1.8908	2.5900	-11.385	9.8796
OLS	15	15	1000	1.9732	1.0847	-1.2642	5.5286
Trim = .2	15	7	854	1.9726	1.3777	-1.9814	6.6415
OLS	20	20	1000	1.9881	0.732	-0.4628	4.0192
Trim = .2	20	10	905	1.9974	0.5242	-0.9019	4.2907
OLS	25	25	1000	1.9920	0.4803	0.0220	3.5389
Trim = .2	25	13	962	1.9821	0.5832	-0.0439	3.7556
Beta (15, 3)							
OLS	10	10	1000	1.9681	1.6534	-2.8753	7.1628
Trim = .2	10	4	638	1.8876	2.1156	-6.2778	10.054
OLS	15	15	1000	1.9970	0.5410	-0.6849	4.8281
Trim = .2	15	7	855	2.0239	1.1505	-1.6284	5.4731
OLS	20	20	1000	2.0036	0.5824	-0.0860	4.0408
Trim = .2	20	10	921	1.9994	0.6974	-0.5005	4.4302
OLS	25	25	1000	2.0017	0.4345	0.4024	3.3919
Trim = .2	25	13	955	1.9987	0.5114	0.3381	3.5111
Normal (0, 625)							
OLS	10	10	1000	2.0818	2.7782	-7.4647	11.293
Trim = .2	10	4	625	2.2094	3.7515	-13.178	17.096
OLS	15	15	1000	2.0248	1.560	-2.3883	7.8513
Trim = .2	15	7	845	2.0113	1.8402	-4.1560	8.2727
OLS	20	20	1000	1.9744	1.063	-1.4266	5.1356
Trim = .2	20	10	942	1.9916	1.149	-2.0927	6.0467
OLS	25	25	1000	1.9803	0.736	-0.0578	4.2606
Trim = .2	25	13	958	1.9847	0.740	-0.5820	4.4788

erated the more accurate mean coefficient estimate in all instances where samples had fewer than twenty observations; these are precisely the cases in which analysts are most concerned about outlier bias. This is probably due to the fact that after trimming, TLS had few observations to work with when data sets were small. An unrelated, but significant, advantage of OLS is that it generated coefficient estimates for every data set. The regression quantiles used for L -estimators are calculated as solutions to a linear programming problem, using the approach developed by Koenker and Bassett. However, for small data sets, there may not be a feasible solution. TLS could not reach a solution in over one-third of the regressions using ten observations, and even when twenty-five observations were available, TLS still failed to reach a solution roughly one-twentieth of the time.

Concluding Comments

This analysis shows that many robust techniques fail to give more reliable coefficient estimates than OLS in a simple detrending regression when an extreme value of the dependent variable lies at the end of a time series. For this class of problem, influential outliers can occur if and only if (a) the standard error of estimate omitting the i th observation is very small and/or (b) the i th residual is very large. However, these conditions rarely occur under probability distributions typical of corn yields. A Monte Carlo study found OLS to generate more accurate coefficient estimates than trimmed least squares. It is concluded that for yield detrending problems, OLS

provides more consistently accurate coefficient estimates than the robust regression alternatives tested. When an extreme outlier is suspected, the *DFBETAS* regression diagnostic statistic provides the best measure of whether its influence is strong enough to merit eliminating that observation from the data set.

[Received March 1990, final revision received August 1990.]

References

- Belsley, David A., Edwin Kuh, and Roy E. Welsch. *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. New York: John Wiley & Sons, 1980.
- Day, Richard H. "Probability Distributions of Field Crop Yields." *J. Farm Econ.* 47(1965):713-41.
- Judge, George G., R. Carter Hill, W. E. Griffiths, Helmut Lutkepohl, and Tsoung-Chao Lee. *Introduction to the Theory and Practice of Econometrics*, 2nd ed. New York: John Wiley & Sons, 1988.
- Koenker, R. W., and G. W. Bassett. "Regression Quantiles." *Econometrica* 50(1978):1577-83.
- Nelson, Carl H., and Paul V. Preckel. "The Conditional Beta Distribution as a Stochastic Production Function." *Amer. J. Agr. Econ.* 71(1989):370-78.
- Swinton, Scott M., and Robert P. King. "Choice of Regression Method for Detrending Time-Series Data with Nonnormal Errors." Dep. Agr. and Appl. Econ. Staff Pap. No. P89-19, University of Minnesota, 1989.
- Weisberg, Sanford. *Applied Linear Regression*, 2nd ed. New York: John Wiley & Sons, 1985.
- White, Kenneth J., Shirley A. Haun, and Nancy G. Horsman. *SHAZAM: The Econometrics Computer Program Version 6*. Dep. Econ., University of British Columbia, 1987.

Cointegration Tests and Spatial Price Linkages in Regional Cattle Markets

Barry K. Goodwin and Ted C. Schroeder

This analysis empirically evaluates spatial linkages in regional cattle markets using cointegration tests of regional price series. Several markets were not cointegrated over the 1980 through 1987 period. However, significant increases in cointegration of several regional livestock markets are observed through the 1980s. The increased cointegration parallels significant structural changes in the livestock industry. A formal analysis of market characteristics reveals that distances between markets, industry concentration ratios, market volumes, and market types have significant influences on cointegration relationships between markets.

Key words: cattle markets, cointegration tests, spatial market linkages.

Spatial price behavior in regional cattle markets is an important indicator of overall market performance. Markets that are not integrated may convey inaccurate price information that might distort producer marketing decisions and contribute to inefficient product movements. In recent years, cattle markets have undergone dramatic regional shifts in production, slaughtering, and processing (Ward). However, some regional markets, particularly those declining in relative volume and importance, may not react efficiently to evolving information (Tomek). Previous studies have examined lead-lag relationships among regional cattle prices (Bailey and Brorsen; Koontz, Garcia, and Hudson; Schroeder and Goodwin). Results generally indicate that prices in the western Corn Belt region tend to lead prices in other markets. With the disparity in market power between cattle producers and slaughter cattle procurers, cattle buyers have the potential to influence regional prices. In addition, with declining relative cattle sales volumes in some regions and increasing volumes in other locations, regional prices could diverge because of poor information flows across regions. In the presence of these influences, price changes across different regions may not fully reflect relevant economic conditions.

The objective of this paper is to empirically evaluate cointegration and spatial price linkages for regional slaughter cattle markets and to determine the impacts on cointegration of several market characteristics. Cointegration tests are developed and applied to spatial price relationships among eleven regional slaughter cattle markets. Bootstrapping regressions are then utilized to evaluate the influences of several variables on cointegration.

Spatial Price Linkages in Regional Markets

Regional markets may be linked through the competitive profit-seeking activities of commodity arbitrageurs. Arbitrage activities commonly involve the purchase of commodities in regional markets where the price is relatively low and the subsequent resale in markets where the commodity is dear. Regional arbitrage will ensure that a unique equilibrium is reached where local prices in alternative markets differ by no more than transportation and transactions costs. In this case, the expected returns to regional commodity price speculation are forced to zero and markets are spatially integrated. However, Takayama and Judge have noted that this sort of direct market linkage requires the assumption that all buyers and sellers are located at discrete points with no intraregional transportation costs.

Alternatively, regional markets may be linked through an oligopolistic (or oligopsonistic) interdependence whereby firms compete only within a limited service area (Benson and Fa-

The authors are assistant professors, Department of Agricultural Economics, Kansas State University.

This is Contribution No. 90-536-J of the Kansas Agricultural Experiment Station.

The helpful comments of Andrew Barkley, David A. Bessler, Merle Faminow, Philip Garcia, and Zach Lea are gratefully acknowledged.

minow). In this case, transportation costs which increase as trade distances rise necessarily limit the area relevant to interfirm competition. According to Faminow and Benson, this situation may result in market linkages that follow a non-competitive basing-point pricing system which is maintained through an organized oligopoly arrangement. Under basing-point pricing, oligopoly cooperation among firms leads to the establishment of a base price at a central location, and the delivered price to any buyer location is the base price plus transportation costs (Scherer). Under a basing-point pricing arrangement, markets will be spatially integrated as the collusive arrangement implies perfectly linked prices at all locations. Faminow and Benson note that this version of spatial integration does not depend upon the assumption that intraregional transportation costs are negligible.

Ardeni has argued that the conventional approaches to testing spatial integration may misrepresent or ignore the time-series properties of regional price data series. Such properties may have important implications for the spatial linkages of a market. In particular, ignoring serial correlation in empirical tests of spatial market integration may result in tests which are vulnerable to inferential inconsistencies and biases. Furthermore, analyses making use of price differentials may suffer because such differencing transformations and filters are ad hoc in nature and may be inappropriate for a given price series.

Alternative procedures for evaluating spatial market linkages have been developed within the framework of cointegration tests by Granger, Engle and Granger, and Engle and Yoo. The general cointegration procedures appeal to the fact that deviations from equilibrium conditions for two economic variables, which are nonstationary when taken by themselves, should be stationary. The intuition is that economic forces should prohibit persistent long-run deviations from equilibrium conditions, although significant short-run deviations may be observed. An important implication is that, while individual economic variables such as prices may wander extensively, certain pairs of such variables should not diverge from one another in the long run. A recent application of cointegration testing to spatial price relationships can be found in Ardeni.

Cointegration tests provide an especially suitable framework in which to consider long-run price relationships among regional cattle markets. Weekly prices in these markets are highly

variable and often possess significant trends, suggesting the potential for nonstationarity in long-run price series. However, efficient arbitrage and basing-point pricing conditions in regional markets suggest that regional prices in alternative markets should not diverge from one another.¹

Cointegration Tests of Spatial Price Relationships

Consider two series of economic variables; y_t and x_t . Each series by itself is nonstationary and requires a single differencing transformation to produce a stationary series. However, a linear combination of the two series,

$$(1) \quad y_t - \alpha - \beta x_t = e_t,$$

produces a residual series e_t which is stationary. In this case, the series y_t and x_t are said to be cointegrated. More precisely, y_t and x_t are cointegrated of order (1,1) with a cointegrating parameter of β and the linear relationship given by (1) is called the cointegrating regression. A series is integrated of order (d) if it must be differenced d times to obtain stationarity. Two series are cointegrated of order (d,b) if the individual series are integrated of order (d) and their linear combination is integrated of order ($d-b$) (Engle and Granger).

Spatial market integration, brought about by commodity arbitrage or collusive basing-point pricing, necessarily implies a unique long-run equilibrium relationship in which deviations from regional price parity are forced to zero. Such an equilibrium need not require that deviations from equilibrium are always zero or even serially uncorrelated. Delivery lags between spatial markets and other impediments to regional trade and arbitrage activities might result in the short-run persistence of deviations from the steady-state equilibrium. Similarly, incomplete basing-point pricing behavior as described by Faminow and Benson might result in short-run deviations from equilibrium conditions. However, spatial integration will require that deviations from eco-

¹ Granger (p. 218) notes that, for a pair of prices from speculative markets, cointegration implies a lack of efficiency because prices should incorporate all available information. However, this result requires instantaneous arbitrage, which is necessarily precluded by spatial linkages that include delivery lags. With such linkages, a spatial dimension to efficiency will require that prices not wander arbitrarily far apart in the long run. In this light, cointegration may imply a form of spatial efficiency for regional market linkages.

nomic equilibrium conditions have a stable mean of zero. In terms of cointegration, this condition implies that two spatially separated markets are spatially linked and interdependent if regional prices can be used to form a cointegrating regression representative of a spatial price parity equilibrium condition.

Conventional analyses of spatial market integration typically posit a parity relationship between markets in which price changes in one market are reflected by equilibrating changes to prices in alternative markets. Such an equilibrium relationship can be written as²

$$(2) \quad p_t^1 - \alpha - \beta p_t^2 = u_t,$$

where p_t^1 and p_t^2 represent logarithmic commodity prices in two alternative regional markets. The residual error term u_t represents proportional deviations from price parity. If the logarithmic price series are nonstationary but their combination in (2) produces a stationary series for u_t , the series are cointegrated of order (1,1) with a cointegration parameter β . This condition suggests a natural empirical approach for evaluating the spatial price linkages between a pair of regional markets through the use of OLS estimation procedures in connection with the cointegration analysis techniques. Perfect spatial integration will be offered empirical support for a pair of regional markets if the regional price series are cointegrated with a cointegrating parameter approximately equal to one.³

The existence of perfectly spatially integrated markets (where price changes in one market are fully reflected by equilibrating changes in alternative markets) necessarily requires that the estimated slope parameter of the cointegrating regression, $\hat{\beta}$, have a value of one. However, because the price series p_t^1 and p_t^2 are nonstationary in a cointegrated system, conventional statistical procedures cannot be used to provide reliable hypothesis tests regarding the value of β . In particular, while the empirical estimate of the cointegrating parameter is consistent, the es-

timated standard error of the cointegrating parameter is not consistent. In this light, it is impossible to conduct explicit significance tests for the estimated cointegrating parameter. Engle and Granger discuss the properties of OLS estimates of cointegrating regression parameters in detail.

Engle and Granger have proposed a two-step procedure for evaluating the cointegrating properties of a concurrent pair of nonstationary economic time series. The first stage involves estimating the parameters of the cointegrating regression representative of the presupposed economic equilibrium condition by using standard OLS regression techniques. Estimates of the parameters of the cointegrating regression can then be used to calculate estimates of the residual errors, \hat{e}_t , where

$$(3) \quad \hat{e}_t = y_t - \hat{\alpha} - \hat{\beta}x_t.$$

Upon obtaining estimates of the first-stage residual errors, Engle and Granger propose seven different tests for cointegration. They also provide critical values for a sample of 100 observations based upon the results of Monte Carlo simulations for each proposed test statistic. Each test has no cointegration as its null hypothesis. Rejections of the tests which follow lend support for cointegration in the regional markets in this analysis. The existence of a variety of tests follows from the usual problems associated with unit root tests⁴ and parameters which are unidentified under the null hypothesis, and because the power of any individual test may vary across actual empirical circumstances (Engle and Granger).

The first test for cointegration suggested by Engle and Granger involves the use of the standard Durbin-Watson test statistic from the first-stage OLS estimate of the cointegrating regression ($CRDW$):

$$(4) \quad CRDW = \left(\sum_{t=2}^T (\hat{e}_t - \hat{e}_{t-1})^2 \right) / \left(\sum_{t=1}^T \hat{e}_t^2 \right).$$

The null hypothesis of no cointegration is rejected for values of $CRDW$ which are significantly different from zero. The $CRDW$ statistic is approximately equal to $2 - 2\hat{\rho}$, where $\hat{\rho}$ is the estimated autoregressive parameter for the residual errors from the cointegrating regression (Judge et al.). In this light, Engle and Granger's

² The parameter α in (2) might be considered a proportional shifter which reflects the costs of interregional trade. Such costs might include interstate tariffs, risk premiums, and transportation charges.

³ While perfect spatial integration requires that price changes in one regional market are fully matched in other markets, cointegration requires only that prices not drift arbitrarily far apart in the long run. We would not necessarily expect efficiently linked markets to have $\beta = 1$. Prices in two efficiently linked markets could vary in a nonsynchronous manner within the band created by transactions costs and thus not imply perfect integration (i.e., $\beta \neq 1$). In addition, quality differences across regions might also bring about a $\beta \neq 1$ without implying inefficient spatial linkages. However, cointegration should be supported in both cases.

⁴ In particular, unit-root test statistics must be constructed using nonstandard distributions, using parameters which are unidentified under the null hypothesis.

first proposed test is analogous to a test of whether $\hat{\rho}$ is significantly different from one.

The second and third cointegration tests proposed by Engle and Granger utilize Dickey-Fuller (1979, 1981) type regressions to consider whether the autoregressive parameter for the estimated residuals from the cointegrating regression is significantly different from one. If there is a unit root, then the two series are not cointegrated. The first Dickey-Fuller-type test depends on estimates of

$$(5) \quad \Delta \hat{e}_t = -\phi \hat{e}_{t-1} + \epsilon_t,$$

where \hat{e}_t is the first-stage estimate of the residual from the cointegrating regression and Δ implies the first difference. A test statistic (DF) is constructed from the ratio of the estimated ϕ to its standard error (a t -ratio). The null hypothesis of no cointegration is rejected for values of ϕ which are significantly different from zero. This is verified by calculated values of DF which exceed the tabulated critical values determined by Engle and Granger. The second Dickey-Fuller type test is analogous to the first except that it contains p lagged values of the differenced residual errors:

$$(6) \quad \Delta \hat{e}_t = -\phi \hat{e}_{t-1} + \theta_1 \Delta \hat{e}_{t-1} + \dots + \theta_p \Delta \hat{e}_{t-p} + \epsilon_t.$$

The addition of the lagged differences ensures that the second-stage residuals of the augmented Dickey-Fuller regression, ϵ_t , are serially uncorrelated. A test statistic (the ADF) is constructed from this regression by again taking the t -ratio for the estimate of ϕ .

The fourth and fifth cointegration tests proposed by Engle and Granger involve the estimation of a vector error correction mechanism for the cointegrating regression:

$$(7) \quad \begin{aligned} \Delta y_t &= \beta_1 \hat{e}_{t-1} + \epsilon_{1t}, \text{ and} \\ \Delta x_t &= \beta_2 \hat{e}_{t-1} + \gamma \Delta y_t + \epsilon_{2t}. \end{aligned}$$

A test of no cointegration is based on the joint significance of the error correction coefficients β_1 and β_2 . This test exploits the fact that a cointegrated set of variables can be equivalently expressed as an error correction model in the form of (7). The intuition behind such error-correcting mechanisms lies in the expectation that agents will react to preceding deviations from equilibrium conditions, thus bringing about responses in the current values of the cointegrated variables, y_t and x_t . A test statistic ($RVAR$) is

calculated by taking the sum of the squared t -ratios of β_1 and β_2 . If β_1 and β_2 are jointly different from zero, the null hypothesis of no cointegration is rejected. The fifth cointegration test statistic (the $ARVAR$) is constructed in an analogous manner to the fourth from an augmented system with added lagged values of differences of the economic variables (Δy_t and Δx_t) to ensure white noise in the error terms of the vector autoregressive system.

The two preceding vector autoregression cointegration tests are conditional on the estimates from the cointegrating regression. The final two cointegration tests relax this restriction by estimating a vector autoregression which is not constrained to satisfy the cointegration constraints. The first system regresses differenced values of each of the hypothesized cointegrated variables on the levels of the variables:

$$(8) \quad \Delta y_t = \theta_1 y_{t-1} + \theta_2 x_{t-1} + c_1 + \epsilon_{1t}, \text{ and}$$

$$\Delta x_t = \theta_3 y_{t-1} + \theta_4 x_{t-1} + \gamma \Delta y_t + c_2 + \epsilon_{2t}.$$

The null hypothesis of no cointegration is rejected if parameters θ_1 and θ_2 of the first equation and θ_3 and θ_4 of the second equation are jointly, significantly different from zero. A failure to reject the null hypothesis indicates the lack of a significant relationship between current changes and past values of the economic variables. It implies a general failure of cointegration between the variables. A convenient test statistic (the $UVAR$) can be constructed by taking twice the sum of the F -tests of joint significance of the θ_i 's in each equation. The final cointegration test statistic (the $AUVAR$) is constructed in an analogous manner from an augmented system with additional p -lags of Δy_t and Δx_t added to ensure serially uncorrelated residuals.

Detailed discussions about these tests are presented in Engle and Granger. Alternative empirical circumstances could alter the power and conclusions of any individual test. Thus, to strengthen the confidence associated with the tests, our application utilizes each of the seven test statistics. In many comparisons of interrelated economic time series, economic theory provides little guidance as to which variable corresponds to x_t and which corresponds to y_t in the cointegration tests. This is certainly true for the bivariate comparisons of alternative regional price series in this study. Because misspecification could bias the results of the cointegration tests, each test is performed twice, with the designations reversed for the second test.

Factors Affecting Cointegration and Spatial Arbitrage Opportunities

The cointegration tests discussed above provide evidence about the linkages among prices at different markets. In this light, cointegration is not an absolute test but is a matter of degree. For a given time period, two markets' price series which move together will be highly cointegrated. Conversely, two markets' price series which tend to diverge from each other for extended periods will have low cointegration. Of interest are the factors that influence the degree of cointegration among regional fed cattle markets. Greater understanding of these factors allows more conclusive statements regarding spatial price relationships among these markets.

A primary factor affecting cointegration is an agent's cost and risk associated with trade between markets (Buccola 1989). The greater are the risks and costs of trade between two markets, the less likely they are to be cointegrated. In general, all markets are linked with one another through intermediary markets. However, these linkages weaken as the distance between markets increases. Mulligan and Fik have formalized this distance relationship in an analytical model of spatial competition. Distance between markets is used as a proxy of costs and as a contributing factor toward risk in this study. The costs associated with trade between spatially separated cattle markets will be directly related to the road distance between the markets. Trucking costs and shrinkage may prevent perfect correlation among prices in spatially separated markets and may even cause prices to diverge for periods of time, even if these markets are operating efficiently. In addition, risks associated with higher death losses and changing relative prices between markets during shipping and delivery lags increase as distance between markets increases. Thus, distance between markets is expected to have a negative influence on the level of cointegration.

A second factor affecting the risk of arbitrage between markets is the amount of market information reflected in prices at a particular market. Liquid terminal markets, for example, have a more complete set of market information in each trade than do decentralized direct trade markets (Lang and Rosa). Price discovery in terminal markets occurs by the interaction of several well-informed packer buyers and commission selling agents. Conversely, direct trades often occur between a single buyer and a single (often less informed) seller. Buccola (1985) finds that cen-

tralized market prices are in general more efficient than decentralized market prices. Thus, with all else constant, terminal markets may be more highly cointegrated than are direct trade markets.

Market volume also may influence trade activity across market areas. Low-volume markets have a higher potential for exhibiting unwarranted price behavior (Tomek). However, the concentration of cattle feeding and alternate market outlets are arguably more important influences on regional trade activities than the absolute volume traded at a particular location. Thus, the concentration of cattle feeding in a market's region could influence the degree of cointegration. However, the influence of market volume on cointegration is not clear. The thin market issue implies that markets with higher volume are more highly cointegrated. However, market regions with higher concentrations of cattle feeding may also operate independently of markets in other regions because of short-term local market forces. Examples are pressures to operate local packing plants at capacity and evolving local information pertinent to price. Thus, the relationship between market volume and cointegration is ambiguous.

A final factor which may influence spatial cointegration in regional cattle markets is packer concentration. Carlton has shown that price stability increases with increasing concentration in several industrial markets. With increased concentration of cattle packers operating across several markets, cointegration of these markets could increase, especially if the highly concentrated firms compete in the same market regions, as occurs in slaughter cattle markets (Ward). By operating plants in several spatially separated markets, the firms reduce spatial transactions costs and uncertainty about market outlets for cattle shipped from one region to another. In addition, increased concentration may facilitate collusive pricing behavior among meatpackers. In particular, discriminatory basing-point pricing practices may result from increased industry concentration. Such behavior would strengthen spatial price linkages.

The seven cointegration test statistics provide formal tests of cointegration which differ with respect to their statistical power under alternative empirical situations. The effects of several market characteristics can be considered through the following model:

$$(9) \quad TS_{ikt}^j = \beta_0 + \beta_1 \text{Type}_i + \beta_2 CR_i + \beta_3 \text{Volume}_{ikt} + \beta_4 \text{Distance}_{ik} + e_{ikt}^j$$

for $i = 1, \dots, 10$; $j = 1, \dots, 7$; and $t = 1, \dots, T$; where TS_{ikt}^j represents the j th cointegration test statistic between markets i and k ($j = 1, \dots, 7$ for $CRDW, \dots, AUVAR$), $Type_i$ is an indicator variable equal to one if market i is direct and zero if a terminal market, CR_t is the four-firm beef slaughter concentration ratio in time t , $Volume_{ikt}$ is slaughter cattle volume in market i 's region relative to that of market k in time t , $Distance_{ik}$ is miles between market i and market k , and e_{ikt}^j is a residual error. Given the preceding discussion, the parameters of equation (9) are hypothesized to have the following signs: $\beta_1 < 0$, $\beta_2 > 0$, β_3 is uncertain, and $\beta_4 < 0$.

Data Description

Weekly price series for Choice, yield grade 2–4, 900–1,100 pound slaughter steers were assembled for eleven U.S. regional markets over the period covering January 1980 through September 1987, yielding a total of 400 weekly observations. The data were collected from the Chicago Mercantile Exchange and from the U.S. Department of Agriculture's *Livestock, Meat, and Wool Market News*. Prices were collected for direct cattle markets of California, Colorado, Illinois, Iowa–Southern Minnesota, Western Kansas, Eastern Nebraska, and Texas Panhandle. Price data were also obtained for terminal markets of Lancaster, Pennsylvania; Omaha, Nebraska; South St. Paul, Minnesota; and Sioux City, Iowa.

These markets include a geographic dispersion of markets which differ in pricing methods (direct versus terminal) and cattle volumes. The markets include those with the largest cattle volumes as well as several small volume markets.

Several of the price series had a small number of missing observations. The total number of missing prices was twenty-six (less than 0.6% of the total data points). The missing prices were proxied by the predicted values from a regression of the series on the 1,100- to 1,300-pound steer price at the same location during the same week.

Empirical Results for Cointegration Tests

Applications of the empirical tests of cointegration are constrained because the relevant critical values have been defined by Engle and Granger

only for samples of 100 observations.⁵ In this light, our applications are to four subsets of the overall data consisting of 100 observations each, which roughly correspond to the two-year periods 1980–1981, 1982–1983, 1984–1985, and 1986–1987. This approach is consistent with the test requirements and allows us to test the impact of changes in industry concentration and relative cattle feeding concentrations across regions. Ward has noted that changes in marketing patterns and industry concentration ratios may have had important structural impacts on the market performance of the beef industry.⁶

The empirical tests of cointegration must be preceded by a test of nonstationarity for the individual economic variables under consideration. Granger suggests using the unit root test developed by Dickey and Fuller (1979, 1981). The test is generated from the following regression:

$$(10) \quad \Delta y_t = \alpha + \phi y_{t-1} + \sum_{i=1}^p \theta_i \Delta y_{t-i} + \epsilon_t,$$

where p is large enough to ensure that the residual ϵ_t is empirical white noise. The Dickey-Fuller test statistic is the ratio of $\hat{\phi}$ to its standard error obtained from OLS regression. The null hypothesis of a unit root is rejected for values of $\hat{\phi}$ which are negative and significantly different from zero using the significance levels calculated by Dickey and Fuller (1981).

The Dickey-Fuller unit root test, based on equation (10), was utilized to test the null hypothesis of a unit root in each of the price series for each period. The test utilized lag orders selected by the minimum value of Akaike's final prediction error (*FPE*). The results supported the presence of a unit root in every case and thus indicate nonstationarity in each of the price series.⁷

Cointegration regressions of the form given by equation (2) were estimated using ordinary

⁵ Engle and Yoo have further investigated properties of the *CRDW*, *DF*, and *ADF* test statistics for alternative sample sizes. Their results provide critical values for the *CRDW*, *DF*, and *ADF* for samples of $n = 200$.

⁶ Ward (pp. 14–15) notes that the four-firm concentration ratio in the steer and heifer slaughtering industry has gone from 39% in 1980 to over 64% in 1986 (see table 3). Increased industry concentration ratios have raised three important concerns. First, increased market power by processors may depress the prices paid for livestock. Second, larger meatpackers may have the market power to drive up retail prices for meat products. Finally, smaller meat processors are concerned that larger competitors will force them out of business (Ward, p. 17).

⁷ The Dickey-Fuller unit root testing results are available from the authors on request.

least squares regression techniques. Given eleven principal markets, 110 different pairwise comparisons are possible. For brevity, considerations of spatial cointegration are limited to integration comparisons between ten of the regional markets and the eastern Nebraska direct, western Kansas direct, and Omaha terminal markets. While these choices are arbitrary, these markets are interesting because of their central location and relatively large marketing volumes. In addition, this consideration allows us to examine differences in spatial price relationships between terminal and direct markets. Only the results for eastern Nebraska are reported. However, any differences between eastern Nebraska and the other markets are highlighted in the discussion.⁸

Cointegrating parameters and standard error estimates for the eastern Nebraska market are presented for each period in table 1. Each comparison is conducted under two specifications. The first specification regresses prices in each of the i regional markets on the eastern Nebraska price. The designation is reversed in the second specification with the eastern Nebraska price being regressed on each of the remaining i market prices. While the estimates of the cointegrating parameters are consistent, the verification of nonstationary price series implies that

the estimated standard errors are not consistent. This condition precludes using these results for formal hypothesis testing in regard to the value of the cointegrating parameters. However, in most cases the cointegrating parameters are close to one in numerical value. Thus, estimates of the cointegrating parameters provide informal support for spatial integration in the regional cattle markets.

The seven cointegration tests were conducted for each specification of the ten market comparisons over each of the four periods. The augmented tests were conducted with $p = 4$ lagged values of Δy_t and Δx_t .⁹ The cointegration test results for the first specification are presented in table 2. In general, the results appear to question the existence of cointegration in the regional cattle markets. Cointegration is supported by 155 of the 280 different cointegration tests for the eastern Nebraska market. As expected, cointegration is strongest for the markets which are relatively close to eastern Nebraska. However, cointegration diminishes as the distance between an individual market and the eastern Nebraska market increases. Cointe-

⁹ Lags are added to the augmented tests to ensure white noise in the residuals. Akaike's *FPE* criterion indicated lag orders of four or less in nearly every case. For the few cases for which a higher lag order was indicated, standard likelihood ratio tests confirmed that four lags were sufficient to ensure whitened residuals. Thus, we follow Engle and Granger and utilize a lag order of four for the augmented cointegration tests.

⁸ Detailed results for the other markets are available from the authors upon request.

Table 1. OLS Estimates of Cointegrating Parameters

Market	Period	β_I	Standard Error	β_{II}	Standard Error
California direct	I	.6419	.0395	1.1369	.0699
	II	.8647	.0284	1.0459	.0344
	III	.9080	.0241	1.0301	.0274
	IV	.9080	.0336	.9707	.0360
Colorado direct	I	.9247	.0333	.9595	.0346
	II	.9166	.0233	1.0258	.0261
	III	.9656	.0188	.9986	.0194
	IV	.9590	.0245	.9800	.0250
Illinois direct	I	.9822	.0234	.9643	.0230
	II	.9983	.0241	.9477	.0229
	III	1.0407	.0176	.9346	.0158
	IV	.9734	.0246	.9669	.0244
Iowa-S. Minn. direct	I	1.0331	.0224	.9254	.0201
	II	.9921	.0207	.9667	.0202
	III	1.0473	.0160	.9335	.0143
	IV	.9685	.0105	.9986	.0186
Lancaster terminal	I	.7435	.0636	.7837	.0670
	II	.7446	.0243	1.2156	.0397
	III	1.0387	.0334	.8741	.0281
	IV	.9923	.0385	.8780	.0341

Table 1. Continued

Market	Period	β_I	Standard Error	β_{II}	Standard Error
Omaha terminal	I	.9968	.0245	.9473	.0232
	II	.9330	.0195	1.0276	.0215
	III	1.0028	.0213	.9552	.0203
	IV	.9430	.0185	1.0218	.0201
St. Paul terminal	I	1.0528	.0277	.8894	.0234
	II	.9718	.0217	.9812	.0219
	III	1.0016	.0171	.9707	.0166
	IV	.9946	.0200	.9673	.0194
Sioux City terminal	I	1.0391	.0291	.8936	.0250
	II	.9905	.0207	.9678	.0203
	III	1.0158	.0199	.9489	.0185
	IV	1.0218	.0212	.9389	.0195
Texas Panhandle direct	I	.8747	.0290	1.0323	.0342
	II	.9169	.0232	1.0262	.0260
	III	.9538	.0219	.9971	.0229
	IV	.9203	.0262	1.0068	.0286
W. Kansas direct	I	.8727	.0326	1.0083	.0376
	II	.9113	.0243	1.0259	.0274
	III	.9286	.0202	1.0289	.0224
	IV	.9290	.0261	.9991	.0281

Note: Each market is compared with eastern Nebraska. β_I refers to estimated cointegrating parameters from Specification I ($p_i = \alpha + \beta_I p_{Nebraska}$) and β_{II} refers to estimated cointegrating parameters from Specification II ($p_{Nebraska} = \alpha + \beta_{II} p_i$).

Table 2. Cointegration Test Results: Specification I ($p_i = \alpha + \beta p_{Nebraska}$)

Market	Test	Test Statistics				Critical Value
		Period I	Period II	Period III	Period IV	
California direct	CRDW	.357	.454*	.469*	.389*	.386
	DF	4.111*	3.894*	3.604*	3.226	3.37
	ADF	2.621	2.605	2.820	2.516	3.17
	RVAR	18.595*	15.531*	12.635	11.144	13.6
	ARVAR	8.242	7.829	7.740	7.540	11.8
	UVAR	19.235*	17.190	14.090	12.846	18.6
	AUVAR	10.563	10.607	11.745	10.198	17.9
Colorado direct	CRDW	.338	.458*	.564*	.636*	.386
	DF	3.840*	3.675*	3.995*	4.307*	3.37
	ADF	3.013	2.335	2.464	3.372*	3.17
	RVAR	14.326*	17.891*	16.595*	20.435*	13.6
	ARVAR	8.863	6.157	6.693	10.933	11.8
	UVAR	14.762	20.160*	18.165	21.756*	18.6
	AUVAR	10.697	10.007	10.569	12.514	17.9
Illinois direct	CRDW	.658*	.501*	.685*	.503*	.386
	DF	4.451*	3.751*	4.470*	3.820*	3.37
	ADF	3.698*	2.116	3.385*	3.058	3.17
	RVAR	29.806*	14.721*	30.208*	27.860*	13.6
	ARVAR	10.989	5.333	10.808	20.474*	11.8
	UVAR	32.526*	16.526	32.658*	31.438*	18.6
	AUVAR	13.845	9.710	16.644	25.518*	17.9
Iowa-S. Minn. direct	CRDW	.493*	.506*	.778*	.685*	.386
	DF	3.726*	3.727*	4.758*	4.641*	3.37
	ADF	2.901	2.200	3.216*	4.063*	3.17
	RVAR	14.078*	14.770*	23.973*	22.328*	13.6
	ARVAR	7.899	5.588	11.833*	23.868*	11.8
	UVAR	14.761	16.628	25.557*	23.826*	18.6
	AUVAR	9.327	8.941	17.938*	26.881*	17.9
Lancaster terminal	CRDW	.243	.922*	.603*	.485*	.386
	DF	2.572	5.388*	4.350*	3.668*	3.37
	ADF	1.973	2.933	2.538	2.837	3.17
	RVAR	19.755*	33.055*	18.234*	26.295*	13.6

Table 2. Continued

Market	Test	Test Statistics				Critical Value
		Period I	Period II	Period III	Period IV	
Omaha terminal	ARVAR	11.756	7.386	6.415	16.387*	11.8
	UVAR	26.503*	34.540*	19.775*	29.556*	18.6
	AUVAR	25.091*	10.413	11.215	21.615*	17.9
	CRDW	.552*	.713*	.405*	.743*	.386
	DF	3.959*	4.611*	3.072	4.829*	3.37
	ADF	2.813	2.052	2.855	3.605*	3.17
	RVAR	21.858*	36.220*	28.602*	42.119*	13.6
	ARVAR	8.277	4.277	12.209*	20.886*	11.8
St. Paul terminal	UVAR	24.013*	38.839*	32.827*	45.062*	18.6
	AUVAR	10.212	9.010	19.591*	24.789*	17.9
	CRDW	.633*	.621*	.761*	.694*	.386
	DF	4.280	4.217*	4.783*	4.726*	3.37
	ADF	2.815	1.985	3.082	3.879*	3.17
	RVAR	21.305*	18.901*	43.104*	38.599*	13.6
	ARVAR	6.649	4.288	15.767*	26.195*	11.8
	UVAR	22.468*	20.773*	45.699*	41.496*	18.6
Sioux City terminal	AUVAR	9.396	7.670	23.784*	31.090*	17.9
	CRDW	.480*	.618*	.677*	.941*	.386
	DF	3.777*	4.239*	4.471*	5.534*	3.37
	ADF	2.420	1.979	2.614	4.302*	3.17
	RVAR	15.490*	24.684*	23.466*	34.830*	13.6
	ARVAR	5.009	3.287	6.822	18.213*	11.8
	UVAR	16.329	26.968*	25.193*	36.573*	18.6
	AUVAR	6.380	6.733	10.834	20.679*	17.9
Texas Pan. direct	CRDW	.395*	.463*	.392*	.439*	.386
	DF	3.888*	3.619*	3.212	3.477*	3.37
	ADF	3.280*	2.790	2.258	3.151	3.17
	RVAR	14.869*	14.661*	11.882	11.864	13.6
	ARVAR	9.736	8.555	6.062	9.198	11.8
	UVAR	15.366	16.516	13.724	13.209	18.6
	AUVAR	11.043	11.622	10.375	10.722	17.9
	CRDW	.390*	.415*	.389*	.400*	.386
W. Kansas direct	DF	3.765*	3.404*	3.244	3.292	3.37
	ADF	2.651	2.770	2.262	2.808	3.17
	RVAR	14.243*	13.241	10.743	12.559	13.6
	ARVAR	7.264	8.192	5.915	7.785	11.8
	UVAR	15.041	15.150	12.324	14.142	18.6
	AUVAR	10.340	11.264	9.917	9.410	17.9

Note: Eastern Nebraska is the central market. Critical values (Engle and Granger) are at the 5% level. An asterisk denotes rejection of the null hypothesis of no cointegration at the 5% level.

gration is strongest between eastern Nebraska and the Omaha, St. Paul, Sioux City, Iowa–Southern Minnesota, Illinois, and Lancaster, Pennsylvania, markets. In each of these cases, seventeen or more of the twenty-eight tests across the four time periods reject the null hypothesis of no cointegration. The Texas Panhandle, California, and western Kansas markets display limited cointegration with the eastern Nebraska market, with ten or less of the twenty-eight tests supporting cointegration. Schroeder and Goodwin and Koontz, Garcia, and Hudson found that price changes in eastern Nebraska and Iowa lead price changes in Illinois, Lancaster, Omaha, St. Paul, Sioux City, and other markets. This result

is consistent with cointegration among these markets.

The western Kansas direct and Omaha terminal markets had results similar to those of eastern Nebraska. Of the 280 cointegration tests, 111 supported cointegration with western Kansas and 176 supported cointegration with Omaha. The western Kansas market was primarily cointegrated with the nearby Colorado and Texas markets. The Omaha market was mostly cointegrated with the same markets as eastern Nebraska.

An important result regarding the behavior of regional market linkages over time is also revealed in table 2. The degree of cointegration

in regional cattle markets appears to have increased over the past several years. In the first two periods (1980–81 and 1982–83), only thirty-four of the seventy cointegration tests support cointegration. However, in the 1984–85 and 1986–87 periods, respectively, thirty-eight and forty-nine of the seventy tests support cointegration. These results suggest that regional market cointegration was enhanced during a time when concentration of the beef industry was increasing.¹⁰ These results were similar although less pronounced for western Kansas and Omaha.

In all, the empirical applications suggest that cointegration of regional cattle prices is limited. This may suggest that markets are segmented and arbitrage opportunities are precluded by high transactions costs or other barriers. Alternatively, the results may indicate the absence of collusive basing-point pricing behavior. However, the results also indicate that structural changes in the beef industry throughout the 1980s were paralleled by increased cointegration of spatial prices in regional direct and terminal markets. This relationship may imply increased pricing efficiency or an increased tendency toward coordinated basing-point pricing by meatpackers.

Faminow and Benson examined the behavior of price differentials in Canadian slaughter hog markets over time to provide additional evidence of basing-point pricing practices. With basing-point pricing, price differentials should reflect transportation costs and thus should increase as the distance between markets increases. Their results indicated that price differentials tended toward equality over time, suggesting the breakdown of basing-point pricing. Absolute price differentials between the eleven markets considered in this study were evaluated for each of the four periods. Although no clear trend toward equality between the price differentials was evident, the range of the price differentials did narrow across time. This result may suggest lower transportation costs and may provide additional evidence in favor of in-

creased pricing efficiency in the regional cattle markets.¹¹

Factors Influencing Spatial Price Linkages

The preceding evaluation of cointegration in spatial markets indicated that regional cattle markets are less than fully cointegrated. In this section, factors which may influence cointegration are formally assessed. A regression-type analysis of equation (9) is undertaken to evaluate the effects of market type, industry concentration ratios, relative slaughter volumes, and spatial distances on cointegration relationships. Table 3 contains the average annual slaughter volumes, distances between markets, and concentration ratios used in this analysis.

Ordinary regression estimates of equation (9) cannot directly be utilized to provide inferences regarding the influences of such factors on cointegration because the test statistic TS_{ita}^j is a generated regressand which follows a nonstandard (nonnormal) distribution. In this light, an alternative estimation strategy using bootstrapping techniques (Efron) is undertaken. Bootstrapping requires only that the residual errors be independently and identically distributed, regardless of their distribution (Prescott and Stengos). Bootstrapped coefficient estimates and standard errors can thus be used to provide consistent inferences about the variables influencing cointegration.

Bootstrapped coefficient estimates and implied *t*-ratios for each of the seven cointegration test statistics are reported in table 4. The bootstrapping estimates were obtained from 1,000 replications. The equations explained from 29% to 54% of the variation in the cointegration test statistics. The coefficient on market type is negative for six of the seven tests (three of which are significantly different from zero at the 5% level), suggesting that direct markets are less likely to exhibit cointegration with the eastern Nebraska market. This result was not consistent for cointegration test statistics obtained from the Kansas and Omaha market comparisons. The coefficient on market type was not statistically different from zero (at the 5% level) for any of the fourteen test statistics for these two markets. Thus, cointegration market linkages do not appear to differ between terminal and direct markets.

¹⁰ Cointegration test results for the alternative specification in which the designations of independent and dependent variables are reversed are not reported here. Although subtle differences exist, the overall results appear to be invariant with respect to the particular specification used in the cointegration tests. Overall, 146 of the 280 efficiency tests supported cointegration in the regional cattle markets. The same patterns were apparent in the results for the alternative market comparisons. In particular, western and southwestern markets exhibited limited spatial cointegration in linkages with eastern Nebraska and cointegration levels appeared to rise over time.

¹¹ Detailed results of the comparisons of spatial price differentials are available from the authors on request.

Table 3. Factors Considered to Influence Cointegration: Distance between Markets, Annual Cattle Slaughter, and Concentration Ratios, 1980–87

Market	Distance from E. Nebraska (miles) ^b	Average Annual Cattle Slaughter ^a (thou. head)			
		1980–81	1982–83	1984–85	1986–87
E. Nebraska direct		5,766	5,355	5,461	5,638
California direct	1,582	1,916	1,736	1,672	1,424
Colorado direct	613	1,589	2,032	1,902	2,019
Illinois direct	394	1,221	1,050	1,246	1,396
Iowa–S. Minn. direct	197	3,184	3,254	2,293	2,051
Lancaster terminal	1,211	739	903	983	1,074
Omaha terminal	58	5,766	5,355	5,461	5,638
St. Paul terminal	289	976	1,101	1,055	1,006
Sioux City terminal	125	3,184	3,254	2,293	2,051
Texas Panhandle direct	590	5,820	5,906	6,461	6,214
W. Kansas direct	380	3,293	4,500	5,774	6,379
Four-firm beef slaughter concentration ratios		40.9	47.8	52.8	64.0 ^c

Sources: USDA, Ward.

^a Average slaughter for the state in which the market is located (Iowa is taken for the Iowa–S. Minnesota market).^b Approximate miles are measured from cities in close proximity to the market under question.^c Includes only 1986, concentration ratio for 1987 was not available.**Table 4. Bootstrapped Coefficient Estimates for Cointegration Test Statistics**

Independent Variable	Dependent Variable						
	CRDW	DF	ADF	RVAR	ARVAR	UVAR	AUVAR
Market type	-0.0924 (-2.20)* ^a	-0.2708 (-1.47)	0.1723 (1.00)	-8.9350 (-4.12)*	-0.6149 (-0.44)	-9.9856 (-4.30)*	-1.9948 (-1.13)
Packer CR ₄	0.0055 (2.43)*	0.0128 (1.27)	0.0296 (3.09)*	0.2822 (2.30)*	0.3812 (4.68)*	0.3065 (2.52)*	0.3828 (3.89)*
Relative slaughter	-0.1812 (-2.95)*	-0.7988 (-2.90)*	-0.4108 (-1.60)	-6.9376 (-2.18)*	-4.3527 (-1.96)*	-7.1615 (-2.11)*	-5.4025 (-2.07)*
Distance ^b	-0.0001 (-3.05)*	-0.0005 (-2.38)*	-0.0004 (-1.98)*	-0.0058 (-2.45)*	-0.0027 (-1.61)	-0.0057 (-2.21)*	-0.0022 (-1.10)
Intercept	0.4918 (3.92)*	4.1960 (7.45)*	1.6280 (3.05)*	19.0730 (2.87)*	-5.5408 (-1.24)	20.4230 (3.04)*	-0.7694 (0.14)
R-squared	.47	.34	.29	.53	.44	.54	.40

^a Numbers in parentheses are *t*-ratios. An asterisk indicates significance at the 5% level.^b Approximate distance from the eastern Nebraska market.

The four-firm packer concentration ratio is highly significant in most cases and is positive in every case for the eastern Nebraska market. Similar results were obtained for western Kansas and Omaha in which twelve of the fourteen coefficients on the packer concentration ratio were positive. This result suggests that increasing

concentration in the beef-slaughter industry may have been paralleled by enhanced spatial market cointegration, and thus greater spatial efficiency of regional cattle markets. Alternatively, this result may imply an increased tendency toward collusive basing-point pricing strategies in regional cattle markets.

Relative slaughter volume in the region had a negative effect on cointegration, suggesting that smaller markets exhibit a greater degree of spatial price dependence than do larger markets. The impacts of relative slaughter volume in the Omaha market were consistent with those of eastern Nebraska in that five of the seven tests had significant, negative coefficients on the relative volume variable. Conversely, none of the relative slaughter volume coefficients were significant for the western Kansas market. Finally, the most consistent result across all markets was that the degree of price cointegration between two separated markets is negatively influenced by the spatial distance between the markets. Twenty of the twenty-one coefficients on the distance variable were negative with fourteen of the twenty significant at the 5% level. This result is consistent with the decaying nature of spatial linkages over increasing distances (Mulligan and Fik).

In all, these results support prior suspicions that spatial market linkages have been enhanced by recent structural developments in the livestock industry. Increased industry concentration has coincided with greater cointegration. Regional market pricing performance also is influenced by distances between markets and slaughter volumes.

Concluding Comments

Overall, cointegration is limited in regional cattle markets. Empirical tests indicate that regional fed cattle prices have not been fully cointegrated during the 1980s. That is, the prices across different market regions have exhibited periods of moderate divergence between one another. Does this suggest market inefficiencies? Or, does this imply the absence of noncompetitive basing-point pricing practices? Finally, why are certain pairs of markets highly cointegrated while others are not? If the last question can be addressed, at least in part, then perhaps something has been learned about spatial fed cattle price relationships.

Markets separated by long distances have lower degrees of cointegration than do markets in close proximity. Markets separated by long distances are most likely linked in an indirect distance-decaying relationship through markets located between them. In this regard, a series of price changes may spread across spatial markets with a degree of lag. Given the high costs and risks associated with transporting fed cattle over long

distances, markets separated by long distances may have prices which diverge from each other for extended periods. This divergence, however, may be warranted by market conditions and may not be large enough to permit profitable trade through regional movements of cattle.

Cointegration has increased over time in the markets analyzed in this study. This increase in cointegration has paralleled increasing concentration in cattle slaughtering. Based on these results, one cannot conclude that increased concentration implies increased cointegration of markets. However, because the largest packers compete for fed cattle in the same market regions, it is reasonable that increased concentration reduces trade and information costs across markets and may contribute to spatial cointegration in the markets in which these firms compete. Moreover, packers may have coordinated price behavior across regions as market concentration has increased.

[Received November 1989; final revision received May 1990.]

References

- Akaike, H. "Fitting Autoregressive Models for Prediction." *Ann. Inst. Statist. Math.* 22(1969):243-47.
- Ardeni, P. G. "Does the Law of One Price Really Hold for Commodity Prices?" *Amer. J. Agr. Econ.* 71(1989): 661-69.
- Bailey, D., and B. W. Brorsen. "Dynamics of Regional Fed Cattle Prices." *West. J. Agr. Econ.* 10(1985):126-33.
- Benson, B. L., and M. D. Faminow. "An Alternative View of Pricing in Retail Food Markets." *Amer. J. Agr. Econ.* 67(1985):296-306.
- Buccola, S. T. "Pricing Efficiency in Agricultural Markets: Issues, Methods, and Results." *West. J. Agr. Econ.* 14(1989):111-21.
- . "Pricing Efficiency in Centralized and Noncentralized Markets." *Amer. J. Agr. Econ.* 67(1985):583-90.
- Carlton, D. W. "The Rigidity of Prices." *Amer. Econ. Rev.* 76(1986):637-58.
- Dickey, D. A., and W. A. Fuller. "Distribution of Estimates for Autoregressive Time Series with a Unit Root." *J. Amer. Statist. Assoc.* 24(1979):427-31.
- . "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root." *Econometrica* 49(1981):1057-72.
- Efron, B. "Bootstrapping Methods: Another Look at the Jackknife." *Ann. Statist.* 7(1979):1-26.
- Engle, R. F., and C. W. J. Granger. "Cointegration and Error Correction: Representation, Estimation, and Testing." *Econometrica* 55(1987):251-76.
- Engle, R. F., and B. S. Yoo. "Forecasting and Testing in

A Comparison of Video Cattle Auction and Regional Market Prices

DeeVon Bailey, Monte C. Peterson, and B. Wade Brorsen

The number of cattle sold through video auctions has increased substantially during the past five years. The nation's largest video cattle auction offered over 450,000 cattle for sale in 1988. This study compares price differences between the nation's largest satellite video cattle auction and three large regional auctions. Prices for both the regional and video auctions were adjusted for quality differences, transportation costs, commissions, and days to delivery. Net prices paid by buyers and received by sellers in video auctions exceeded the prices for three major regional auction markets.

Key words: cattle, regional auctions, transactions costs, video auction.

Video auctions are selling increasingly large numbers of cattle. Superior Livestock Auction (SLA), the nation's largest satellite video cattle auction, sold more than 270,000 cattle in 1987 and more than 450,000 in 1988. The SLA is projected to be the largest cattle auction of any kind by 1990 (Scharlier).

Acceptance of any pricing method depends on the motivations of buyers and sellers (Buccola). Sellers will want to use video cattle auctions if they provide relatively high prices and a reliable market. Economic theory suggests that video auction prices may indeed be higher than traditional auction prices. Satellite video auctions reduce travel time and expense for buyers who can bid from remote locations. Video cattle auctions can also reduce buyers' search time since they can offer a large number of cattle quickly.¹ For example, SLA offered over 90,000 cattle for sale during a two-day auction in 1988 (Scharlier). Using video auctions reduces health problems because cattle are not mingled with those from other lots and are transported to only one destination.

Despite these benefits, many buyers and sellers, as well as institutions, are concerned about the economic viability of the video auction as opposed to that of more traditional auctions. For example, because of a state policy requiring economic need or benefit to be demonstrated before licensing new cattle auctions, North Dakota refused a business license to SLA in 1986. At that time no statistical evidence was available to document the benefits of video cattle auctions. In addition, buyers and sellers are concerned about prices of the cattle sold through video auctions relative to prices in other markets.

The effect of electronic marketing systems' on structure and relative prices has been studied extensively (Sporleder; Sporleder and Mahoney; Ward; Rhodus, Baldwin, and Henderson and others). However, no study has examined relative prices of satellite video cattle auctions and traditional auctions.

Visual information provided by video auctions may be superior to the information about standard grades and quality factors provided by electronic markets. This may partly explain the general failure of electronic livestock markets while electronic marketing of alfalfa hay and cotton has been successful. Also, because video auctions operate similarly to traditional livestock auctions, buyers may feel more comfortable with them than with other electronic markets.

This paper compares the prices received for cattle sold through video auctions with prices received for cattle sold in three large traditional regional auctions. It also provides previously

The authors are an associate professor and a research assistant, respectively, Department of Economics, Utah State University; and an associate professor, Department of Agricultural Economics, Purdue University.

Utah Agricultural Experiment Station Journal Paper No. 3859. Financial support was provided by the Research Institute on Livestock Pricing, Virginia Polytechnic Institute and State University, and the Utah Agricultural Experiment Station.

The authors give special thanks to Terrence F. Glover and Donald L. Snyder for helpful comments on an earlier draft.

¹ Other procurement costs include transportation costs, commissions to order buyers, travel costs, subscription fees, salaries, and other search costs.

unavailable descriptive information about video cattle auctions.

Description of Video Cattle Auctions

Video auction cattle presentations consist of two components—the video or visual component and the sales catalogue or written component. A \$2.00 per head videotaping fee is included in the sales commission unless the seller rejects the bid, in which case the seller forfeits the taping fee. The taping is done by one of SLA's regional representatives. Thus, the integrity of the video auction is heavily dependent on the integrity of its regional representatives. Sales catalogue descriptions are prepared by the video auction company and the seller when the cattle are videotaped.

Videotapes of about two-minute duration are shown while an auctioneer solicits bids. Buyers must register in advance of the sale and undergo a credit check in order to participate. Buyers may bid either in person or by telephone from any location where the satellite transmission can be received. SLA used two separate central locations (Denver and Ft. Worth) for its auctions in 1987, with most fall sales held in Denver and all others in Ft. Worth.

The video auction representative oversees delivery. Completed sales become cash forward contracts since all cattle are sold for future delivery. Although the video auction representative is responsible for ensuring contract compliance by both buyer and seller, buyers are permitted to be present at delivery and may attempt to influence which cattle are actually delivered. This may be a problem if prices decline before delivery. SLA believes that most larger volume buyers hedge their cattle purchases, but smaller buyers may not. Consequently, most of the few disputes that have arisen at delivery are with small volume purchasers (J. Odle, Superior Livestock Auction, personal communication, Jan. 1990).

Reasons for Price Differences between Market Types

Markets for specific agricultural commodities differ by space, structure, time of delivery, and quality (Buccola, Schroeder et al., Janssen and Shane, Bailey and Brorsen, Faminow and Gum, and others). Differences examined here are due to the markets, not to space, quality, or time of delivery.

Several differences between video and regional auctions suggest video auctions may sell cattle at prices higher than regional auction prices. The most obvious difference affecting prices is in transactions costs (transportation, shrink, and commissions). If transactions costs are low, sellers could actually receive a relatively high net price in a particular market at the same time buyers pay a relatively lower net price in the same market. If lowered transactions costs reduce their net prices, buyers may bid higher at video auctions than they would at regional auctions.

Cattle sold through video auctions do not need to be shipped to a central location as they do for regional auctions. This saves on both trucking costs and shrink. Video auctions may also have a more competitive market structure than traditional auctions have. More buyers and sellers usually participate in SLA auctions than in the regional auctions, which is indicative of the large area served by SLA and its relatively large sales volume (table 1). Hamm, Purcell, and Hudson found anonymous bidding such as the video auction uses tends to increase cattle prices. Larger sales volumes and anonymous bidding suggest prices at the video auction may be higher than traditional auction prices.

Feedlot operators often perceive cattle sold through traditional auctions as being "severely stressed and breaking with disease immediately upon or soon after arrival" at the feedlot (Pollreis, p. 24). Cattle bought directly from the farm tend to have fewer health problems than cattle sold at regional auctions (Pate). Buyers should be willing to pay higher prices for healthy cattle than for severely stressed cattle.

Video auction catalogues provide buyers with important information usually unavailable in traditional auctions, including vaccinations received, place of origin, and current feed (see fig. 1). This information could reduce buyers' death losses, reduce veterinary costs, and increase feedlot efficiency.

Video and Regional Auction Data Collection and Analysis

The 1987 sales catalogue data and accepted bid prices for each lot of cattle were gathered from the SLA. In 1987, more than 335,000 head were offered for sale via fourteen satellite video cattle auctions (table 2). SLA prices were compared to prices at three regional markets.

The Oklahoma City, Oklahoma (OKC); Greeley, Colorado; and Dodge City, Kansas, re-

Table 1. Average Number of Buyers and Estimated Commissions for Regional and SLA Auctions, 1987

	Day of the Week Most Sales Held	Average Number of Buyers Viewing Auction	Major Buyers Attending ^a	Sales Commissions and Other Deductions For Yearling Steers
SLA ^b	Saturday	225 ^c	30	2% of gross sales + 1.50/head ^d
OKC ^c	Monday	30	15	\$7.34/head
Greeley ^f	Tuesday	50	15	2% of gross sales + 1.50/head
Dodge City ^g	Wednesday	50	20	\$7.20/head

^a Buyers who frequently buy relatively large numbers of cattle.

^b Estimates provided by SLA.

^c Average number of registered buyers with SLA. Of this number, 60–80 will actually buy cattle at an average sale. During 1988, 1,507 sellers consigned cattle to SLA and 372 different buyers purchased cattle.

^d The \$1.50 per head is estimated cost of beef board deduction and inspection.

^e Estimates provided by Oklahoma National Stockyard Company for yearling steers, (personal communication, May 1989).

^f Estimates provided by Greeley Producer Livestock Auction, Greeley CO (personal communication, May 1989).

^g Estimates provided by Dodge City Winter Feeder Cattle Auction, Dodge City KS (personal communication, Feb. 1990).

gional auction markets were selected for price comparisons with SLA based on their sales volume, proximity to high-density feeder cattle populations, and the availability of reliable price data. As shown in table 1, the highest volume sales days for the OKC, Greeley, and Dodge City auctions were Monday, Tuesday, and Wednesday, respectively. Most SLA auctions were held on Saturdays in 1987.

U.S. Department of Agriculture (USDA) personnel report price ranges by sex for 100-pound increments at the OKC, Greeley, and Dodge City markets. The reporters estimate the price range for a given weight and grade class by attempting to identify a price pattern within that class. If lots do not fit the reporter's established subjective pattern for the grades (i.e., No. 1 or No. 2

steers), then the prices for these cattle are either reported separately as a "package" or not reported. Bid prices for cattle that are sick, over filled, have physical defects or excessive hair or dirt, dairy or other "non-typical" breeds are also excluded by USDA reporters when price ranges are established (USDA AMS, Market News Div., Greeley CO and Oklahoma City OK, personal communications, 1989). SLA's sales catalogue reported no cattle with physical defects or excessive hair or dirt. All lots of dairy cattle were eliminated from the SLA data to maintain consistency between SLA's reported prices and those reported by USDA for the regional auctions.

The midpoints of USDA's reported price ranges at the regional auctions on their highest sales volume day of the week are compared to

LOT #1410

75 STEERS

LOCATION: Woodruff, UT—393 miles W of Cheyenne, WY.

DESCRIPTION: Herefords. BWF. Blacks and a few BeefMaster cross.

ORIGIN: Native

FRAME: Medium

EST. WT. VAR.: Uneven

FEED: Pasture with no supplement of any kind.

DELIVERY DATE: Oct. 1–10, 1987. Seller's option.

WEIGHING COND: Cattle will be gathered into dry lot by 8:00 A.M., hauled 15 miles. Unloaded and weighed in Ross Jackson scales in Randolph, UT. on the ground with a 2% pencil shrink.

SLIDE: .03 over 725 lbs.

COMMENTS: Steers have not been implanted. A nice set of light yearling steers.

REPRESENTED BY: Jerry Goodwin

Arlo Eastman

BASE WT:700#

FLESH: Medium to Light

HORNS: None

PRICE _____ BUYER _____

Figure 1. Sample of SLA catalogue entry

Table 2. Selected Characteristics of SLA Video Cattle Auctions for All Lots, 1987

Characteristic	Total
Total number of lots	2,222
Total number of cattle offered for sale	335,654
Total number of cattle sold	271,079
Total number of light steers offered for sale (under 600 lbs.)	68,070
Total number of light feeder steers sold	57,851
Total number of heavy feeder steers offered for sale (600–800 lbs.)	91,624
Total number of heavy steers sold	75,855
Average lot size (head)	150
Average estimated weight per head (lbs.)	613
Average days to delivery	39
Average miles shipped after sale	471

the bid prices actually accepted by sellers at SLA (USDA AMS 1987). For example, price comparisons for steers sold through SLA weighing between 600 and 699 pounds were made against the midpoint price for 600- to 700-pound steers at the three regional markets during the week following each SLA auction, while video auctioned steers weighing from 700 to 799 pounds were compared to the reported regional midpoints for 700- to 800-pound steers.²

Method for Expressing Seller and Buyer Prices

Transactions costs such as trucking and shrink, commissions (usually paid by the seller), and other price and/or weight adjustments, such as pencil shrink, cause seller's and buyer's prices to differ.³ Consequently, prices are examined from both the seller's and buyer's point of view. Since costs are not equal across markets, net

prices are derived and used for comparisons. The expected net price per hundredweight received by sellers and paid by buyers in any particular market can be expressed by the following equations:

$$(1) \quad SNP_i = BP_i(1 - PS) - STC - C,$$

$$(2) \quad BNP_i = BP_i(1 - PS) + BTC,$$

where SNP_i and BNP_i are the seller's and buyer's net prices, respectively, for the i th lot of cattle. BP_i is the successful bid or buy price offered by the buyer to the seller, PS is the percentage pencil shrink in decimal form or any other systematic adjustment to price in a particular market. STC and BTC are transportation costs (trucking and shrink) for the seller and buyer, respectively, and C is commissions.

Video auction cattle sold in 1987 were delivered from a few days up to nine months after sale. Feeder cattle futures prices were used to adjust buyers' and sellers' net prices for different delivery dates. Futures market prices serve as a standard of price comparison in the feeder cattle industry (Just and Rauser), and cash prices adjust quickly to changes in futures prices (Oelermann et al.). The prices are adjusted using the spread between the nearby feeder cattle futures contract and the futures contract for the contract closest to but not preceding the SLA delivery date. Only prices for steers sold through the video auction weighing between 600–800 pounds were used in order to coincide weight specifications for the feeder cattle futures contract.

The seller's and buyer's net price after adjustment for transactions costs, quality, and delivery dates for cattle sold through the video auction can be expressed as

$$(3) \quad ASNP_{it} = SNP_{it} + [NBF_{t+j} - \{(FP_{f,t-1} + FP_{f,t+2})/2\}] - QD,$$

$$(4) \quad ABNP_{it} = BNP_{it} + [NBF_{t+j} - \{(FP_{f,t-1} + FP_{f,t+2})/2\}] - QD,$$

where $ASNP_{it}$ and $ABNP_{it}$ are the seller's and buyer's adjusted net prices in dollars per hundredweight for the i th lot of cattle sold at time t . NBF_{t+j} is the closing quote for the feeder cattle futures contract closest to maturity j days after each video auction where j corresponded to the number of days between the SLA and regional auction sales (i.e., $j = 2$ for OKC, $j = 3$ for Greeley, and $j = 4$ for Dodge City). $FP_{f,t-1}$ and $FP_{f,t+2}$ were the closing quotes for the preceding Friday and following Monday feeder cattle futures contract for the contract closest to but

² An examination of the reported price ranges for the three regional auctions for the weeks preceding and following each SLA sale revealed these ranges to be relatively narrow (the narrowest range was \$3.88/cwt and the widest range was \$6.38/cwt during 1987). The Dodge City data from Kansas State University revealed, after eliminating sick, physically defective, and dairy cattle, that a range of about \$10/cwt existed for the same weight categories.

³ Pencil shrink refers to a systematic downward weight adjustment on cattle designed to make weights at the seller's location approximately equivalent with regional auction weights.

not preceding the delivery date (contract f), respectively. Friday's and Monday's futures prices were used since all but one of SLA's sales were held on Saturday when no futures quote was available. *QD* represented an adjustment for any quality differences that may exist between the two types of markets.

Adjustment for Transactions Costs

The transportation costs and commissions defined in equations (1) and (2) are not equal across markets. The distance cattle were shipped after the video auction sale was obtained from SLA. Distances for the seller's location to the regional auctions and from the regional auctions to the location designated by the buyer are on SLA's shipping records. Buyer's and seller's transportation costs were estimated for cattle sold at the video auction. Transportation costs that would have been incurred by buyers and sellers had the cattle been sold at regional auctions instead were also estimated.

Buyers' and sellers' prices in OKC, Greeley, and Dodge City are compared only with prices of cattle sold through video auction that were within the assumed market areas of the regional auctions.⁴

Transportation costs include both trucking fees and shrink. Potential trucking fees were estimated as the average rates of a large western trucking firm in 1987, which were \$2.00 per loaded mile for distances of less than 150 miles and \$1.85 per loaded mile for distances exceeding 150 miles (Chuck Webb, L. W. Miller Transportation, Logan UT, personal communication, Nov. 1988). Shrink was estimated by the method given in Minish and Fox. Their shrink equation is as follows:

$$(5) \quad \text{Shrink \%} = 3.0 + .0061 (\text{MILES}),$$

where *Shrink %* is the percentage shrink, and *MILES* is the distance cattle are trucked. Seller's potential shrink losses from shipping to a regional auction were calculated as the product of the regional midpoint price for the appropriate weight range and equation (5) using the mileage (*MILES*) from the seller's location to the regional market. This figure was adjusted downward by 25% to account for reductions in

shrink when cattle are allowed to eat and/or drink after arrival at the auction and before sale (Minish and Fox).

Buyer shrink costs for cattle purchased through the video auction were calculated as the product of the video auction bid price adjusted for pencil shrink and equation (5) using the distance from the seller's location to the shipment destination designated by the buyer. This figure was then adjusted downward by 50% since excretory shrink will be gained back quickly after arrival (Minish and Fox). Potential buyer shrink costs had the cattle been purchased at a regional auction were also calculated as the product of the regional auction midpoint prices and equation (5) using the distance from the regional auction to the shipping destination selected by the buyer for each SLA lot. The shrink rate of cattle from the regional auction is expected to be only about 60% of the rate of cattle shipped to the regional auction by the seller (Norris J. Stenquist, Utah State University, personal communications, Feb. 1990). Therefore, estimated buyer shrink costs in equation (5) were adjusted downward by 70% for shipping cattle from the regional auctions. Regional auction prices (*BNP*) are adjusted downward by the percentage shrink that would have been incurred by the seller in shipment to the regional auctions. This adjustment accounts for higher prices being paid for "pre-shrunk" cattle at the regional auctions and makes prices for current delivery at the two types of markets approximately equivalent except for transactions costs differences.

Since the buyer pays transportation costs in the video auction, video auction sellers' net prices [equation (1)] need not be adjusted for shrink and trucking costs (i.e., $TC = 0$). However, the seller usually offers pencil shrink in a video auction ($PS > 0$) to make weights approximately comparable to weights at traditional auctions. Seller's net prices at regional auctions were adjusted for the potential transportation costs from the seller's location to the regional auction.

Adjustments for Quality

Concerns about possible quality differences between the markets prompted an investigation of average quality characteristics of cattle sold in the video and traditional markets. No information was available for individual lots sold in either Oklahoma City or Greeley. However, data assembled by Kansas State University provided individual lot characteristics of cattle sold at Dodge City during March and April of 1987

⁴ An industry expert estimates cattle will usually not spend more than eight hours in transit to a regional auction (Steve Reed, Western Livestock Marketing Project, personal communication, Dec. 1988). Also, comparison of prices for lots outside of the market areas yielded basically the same results that are reported in this paper. However, the variability in prices was greater.

(Schroeder et al.). An industry expert stated that there was no reason to believe that lot characteristics should vary substantially between the Dodge City and the OKC auctions (Michael Sands, Western Livestock Marketing Project, personal communication, Nov. 1988). Supporting this conclusion, average lot sizes were quite similar for the two auctions, with OKC and Dodge City averaging thirteen and twenty head per lot, respectively (USDA AMS Mkt. News Div., Greeley CO, personal communication, Nov. 1989; Schroeder et al.). The Dodge City data represented all lots sold through that auction during these months. To be consistent with both the SLA data and USDA's reporting methods for the other auctions (OKC and Greeley), Dodge City lots with physically defective or sick cattle, dairy cattle, and those with excessive mud or hair were eliminated from the study.

The process for determining the impact of quality on auction cattle prices in a particular market is well documented in the literature. A hedonic specification such as this regresses prices on quality and market characteristics in a form similar to that used by Schroeder et al. and given below

$$(6) \quad P_i = a + \sum_{k=1}^K h_k LC_{ki} + \sum_{l=1}^L g_l MC_{li} + e_i,$$

where P_i is the price of the i th lot of cattle in dollars per hundredweight, LC_k is the k th lot characteristic, MC_l is the l th market condition, a is the intercept, h and g are parameter estimates, and e is an error term.

The parameters of equation (6) were estimated for lots of steers weighing between 600 and 800 pounds that were sold through the video auction and represent the money values of differences in the individual quality and market characteristics. The vector of parameter estimates was then multiplied by the vector of average observed values for the physical attributes and market conditions of the steers sold by SLA and those sold in Dodge City. The difference between this average predicted price provided a measure of the relative difference in the price per hundredweight (cwt.) of cattle sold in the two markets (SLA and Dodge City) based on quality. This differential was also assumed for the OKC and Greeley markets.⁵

⁵ This procedure assumes the hedonic price structure is the same for all markets. The parameters of equation (6) were also estimated for the Dodge City data. The signs and magnitudes of the coefficients were basically the same as those obtained using the SLA data with the exception of lot size.

The average quality differential between the Dodge City and SLA lots was \$0.34 per hundredweight. This figure (\$0.34/cwt.) was subtracted from the video auction price to compare prices for cattle of equal quality at both the video and three regional auctions [QD in equations (3) and (4)].

Tests

A paired t -test (Chou, p. 334) was used to test for significant differences between the seller's adjusted net price at the video auction [$ASNP$ in equation (3)] and the net price that would have been received had the cattle been shipped for current delivery to a regional auction [SNP in equation (1)]. The tests were performed for cattle sold through SLA and lying within each regional market area.

A paired t -test was also performed to determine if the adjusted buyer's net price [$ABNP$ in equation (4)] was significantly different from the estimated net price the buyer would have paid had the cattle been purchased at one of the three regional auctions instead [BNP in equation (2)] and shipped from the regional auction to the destination specified by the buyer on SLA's shipment records.

Price trends were not expected to have much effect on prices separated by only three days. Also, the adjustment using futures prices in equations (3) and (4) should remove any such possible effect. Nevertheless, we conducted a fragility test of the results by comparing the regional auction prices for the week previous to the SLA auction. The overall results were unchanged. However, average regional auction prices were slightly lower for the week previous to our comparison, indicating a positive price trend. A further investigation of price trends using daily futures prices showed that prices trended upward about \$0.03 per hundredweight per day during 1987, not enough to affect the results.

Results

Table 3 gives the parameter estimates for equation (6) and video and Dodge City auction average quality characteristics of 600- to 800-pound steers. The steers sold by the video auction were slightly heavier than those sold in the Dodge City auction. Breeds were more highly aggregated for the video sales than for Dodge City sales, and 67% of video lots were predominantly English-

Table 3. Parameter Estimates for Hedonic Pricing Model for Video Auction [Equation (6)] and Average Quality Characteristics for 600–800 lb. Steers at the Dodge City and SLA Auctions

Independent Variable	Parameter Estimate	Dodge City	SLA
Intercept	24.665 (7.31)** ^a		
Nearby futures price (\$/cwt)	0.931 (30.37)**	69.43	69.74
Number in lot (head)	0.004 (1.84)*	20.9 ^b	180.3
Number squared	−0.000 (1.04)	1,051.99 ^b	60,604.1
Average weight per head (lbs.)	−0.026 (−10.05)**	699.4 ^b	718.2
Not uniform lots ^b	−0.107 (−.37)	.02 ^h	.317
Horns ^c	−1.611 (−1.20)	.076 ^h	.009
Mixed horns ^d	−0.42	.203	.160
<u>Predominant Breed:^e</u>			
English-cross	0.523 (0.381)	.287 ^b	.145
Exotic-cross	−1.063 (−0.77)	.166 ^b	.171
English-exotic cross	0.234 (0.17)	.346 ^b	.670
Angus	−3.841 (−1.29)	.051 ^b	.002
<u>Frame:^f</u>			
Large	4.036 (3.57)**	.051 ^b	.071
Medium-Large	2.141 (2.09)	.880 ^b	.438
Medium	1.611 (1.58)	.070 ^b	.469
<u>Flesh:^g</u>			
Heavy	−4.756 (−2.25)**	.010	.005
Medium Heavy	−4.225 (−4.15)**	.130	.162
Medium	−4.147 (−4.24)**	.840	.719
Light-Medium	−4.444 (−4.20)**	.010	.090

^a t-values are in parentheses; single asterisk denotes significantly different than zero at 10% confidence level, double asterisk denotes significantly different than zero at 5% confidence level or better.

^b Binary variable equal to 1 if a nonuniform lot, zero otherwise.

^c Binary variable equal to one if entire lot is horned and zero otherwise.

^d Binary variable equal to one if some cattle but not all in lot had horns and zero otherwise.

^e Binary variables for breed characteristics. Base was Herefords.

^f Binary variables for frame characteristics. Base was small-framed.

^g Binary variable for flesh characteristics. Base was light fleshed.

^h Denotes a significant difference between the average value for the regional and video auctions at the 5% confidence level.

exotic crosses compared to only 35% of the lots sold at Dodge City. This resulted in a small positive quality differential for the video compared to the regional auction cattle lots since English-exotic crosses received a premium over the base

breed, Herefords. Average values for frame and flesh are similar in the two markets.

A large portion of the difference in quality between the auctions resulted from differences in lot size. When a truckload of reasonably uni-

form cattle is shipped to a regional auction, they are typically sorted into smaller lots before being sold. While the average video auction lot size is 150, the largest lot sold in the Dodge City market was 116 head. Some arguments could be made that no adjustment should be made for the larger lot sizes in video auctions since the regional auction markets have no means of selling large lot sizes. The conservative approach is taken here by adjusting for lot size.

Table 4 presents the unadjusted and the estimated adjusted prices paid by buyers and received by sellers in the different markets. SLA's unadjusted prices were significantly lower than OKC and Dodge City prices during 1987 but were basically equal to Greeley prices. However, after accounting for differences in transportation costs, quality, and different delivery dates, the seller's adjusted net price was significantly higher at the video auction than at the three regional markets (table 4). The average differences among SLA's adjusted price and the adjusted prices at the other three markets were \$0.95 per hundredweight, \$3.36 per hundredweight, and \$1.48 per hundredweight at OKC, Greeley, and Dodge City, respectively. This means that, on the average, revenue from a 700-

pound steer would be between \$6.65 per head and \$23.52 per head higher through the video auction than at the traditional regional auctions.

Buyers appear willing to pay significantly higher prices for cattle purchased at video auctions than for those purchased at traditional auctions. This difference is generally smaller than the difference in sellers' prices and was \$0.66 per hundredweight at the OKC market, \$2.41 per hundredweight at the Greeley market, and \$1.69 per hundredweight at the Dodge City market. This indicates that buyers are willing to pay between \$4.62 per head and \$16.87 per head more for a 700-pound steer purchased through a video auction than for a steer purchased at a traditional regional auction.

Tables 5 and 6 report the average adjustments made to prices. These adjustments include shrink, trucking, commissions, quality, and delivery dates. The major reason sellers receive higher prices in the video auction than in traditional auctions is lower transactions costs [line (e) of table 5]. For example, adjusted seller's prices are \$1.48 per hundredweight higher for video auction cattle than for Dodge City auction cattle, while seller's transactions costs are \$1.25 per hundredweight lower for video auction cat-

Table 4. Average Adjusted and Unadjusted Prices for Video and Regional Cattle Auctions for Sellers and Buyers

Variable	Auction Market Area		
	OKC	Greeley	Dodge City
Observations	200	87	148
Average unadjusted prices:		(\$/cwt.)	
(a) Regional auctions	72.55	72.20	72.10
(b) Video auctions	70.46	72.53	71.29
(c) Test for difference in prices ^a [(b) - (a)]	-12.36***	1.01	-4.52**
Average adjusted sellers' prices:			
(d) Regional auction	67.66	66.82	67.70
(e) Video auction	68.61	70.14	69.19
(f) Difference in prices [(e) - (d)]	0.95 (2.33) ^c	3.36 (1.87)	1.48 (1.97)
(g) Test for difference in prices ^a	5.0**	10.67**	9.13**
Average adjusted buyers' prices:			
(h) Regional auction	72.90	72.13	72.05
(i) Video auction	73.54	74.54	73.68
(j) Difference in prices [(i) - (h)]	0.66 (2.20)	2.41 (3.03)	1.69 (2.27)
(k) Test of difference ^a	4.07**	7.42**	8.79**

^a *t*-statistic from a paired difference test.

^b Asterisks denote cases where the null hypothesis of no difference in means is rejected at the 5% level.

^c Standard deviations are in parentheses.

Table 5. Sellers' Estimated Average Transaction Costs and Adjustments for Quality and Delivery Dates for Video and Regional Auctions

Variable	Auction			
	SLA	OKC	Greeley	Dodge City
	(\$/cwt.)			
(a) Pencil shrink	1.51	N/A	N/A	N/A
(b) Commissions and fees	1.64	1.03	1.66	1.02
(c) Trucking	N/A	1.24	1.18	0.99
(d) Shrink ^a	N/A	2.61	2.55	2.39
(e) Total transaction costs ^b [(a) + (b) + (c) + (d)]	3.15	4.88	5.39	4.40
(f) Transactions cost as a percentage of unadjusted prices	4.4	6.1	7.5	6.7
(g) Adjustment for quality ^c	0.34	N/A	N/A	N/A
(h) Average adjustment for different ^d delivery dates	N/A	1.65	1.09	1.39

^a Shrink also represents the downward adjustment to the adjusted buyers' net price [equation (2)] to account for the preshrink of cattle sold at regional auctions.

^b Transaction costs are subtracted from sellers' prices [equation (1)].

^c The adjustment for quality is subtracted from the sellers' price for the video auction.

^d The adjustment for delivery dates is added to the video auction prices [equation (3)].

Table 6. Transportation Costs and Other Adjustments to Buyers' Bid Prices

Variable	Auction Market Area		
	OKC	Greeley	Dodge City
	(\$/cwt.)		
Average transportation costs:			
(a) Video auction ^a	3.21	2.72	2.79
(b) Regional auction ^b	2.96	2.47	2.28
Transportation costs as a percentage of unadjusted prices:			
(c) Video auction	4.6	3.8	3.9
(d) Regional auction	4.1	3.4	3.2
Average adjustments to prices for shrink:			
(e) Video auction ^c	1.49	1.48	1.49
(f) Regional auction ^d	2.60	2.54	2.39
Average adjustment to prices for delivery dates:			
(g) Video auction ^e	1.65	1.09	1.39
(h) Regional auction	N/A	N/A	N/A
Adjustment to prices for quality:			
(i) Video auction ^f	0.34	0.34	0.34
(j) Regional auction	N/A	N/A	N/A
Total average adjustment to buyers' bid prices:			
(k) Video auction ^g [(a) - (e) + (g) - (i)]	3.03	1.99	2.35
(l) Regional auction ^h [(b) - (f)]	0.36	-0.10	-0.11
Combined buyer and seller transactions costs as a percentage of unadjusted prices:			
(m) Video auction ⁱ	9.0	8.1	8.3
(n) Regional auction ^j	10.8	10.9	9.4

^a Trucking and estimated shrink costs from sellers' location to destination specified by buyer on SLA's shipping records.

^b Trucking and estimated shrink from regional auction to destination on SLA's shipping records.

^c Pencil shrink specified in SLA's sales catalogues.

^d Estimated shrink incurred by seller if cattle had been shipped to the regional auction. This estimates the value to a buyer at a regional auction of cattle that have already experienced some shrinking.

^e Average spread basis between the nearby feeder cattle futures contract and the contract closest to SLA's deliveries.

^f Estimated quality differential between SLA and regional markets.

^g Subtracting this figure from the unadjusted video bid price yields (approximately) buyers' adjusted video price.

^h Subtracting this figure from the unadjusted regional auction bid prices yields (approximately) buyers' adjusted regional auction price.

ⁱ Seller average transactions costs at SLA \$3.15/cwt. (table 5) plus buyer transactions costs [line (a)] divided by unadjusted SLA bid price [line (b) of table 4].

^j Line (b) plus line (c) of table 5 divided by line (a) of table 4.

tle than for Dodge City auction cattle. Unadjusted video prices also reflected relatively large discounts for future delivery [line (h) in table 5].

Sellers pay about 6.1% of the value of their cattle to market them in Dodge City, 7.5% in Greeley, and 6.7% in OKC. However, sellers' transactions costs are only about 4.4% of the value of the cattle at SLA, since sellers pay no trucking in the video auction (table 5).

Buyers' transportation costs are slightly lower from the regional auctions than from the video auction, since cattle are "pre-shrunk" (table 6). Transactions costs are substantial for both sellers and buyers at regional auctions. Combined buyer and seller transactions costs for the video auction average about 8.5% of the average unadjusted bid. Combined estimated buyer and seller transactions costs average approximately 10.4% of the unadjusted bid price for the regional auctions (table 6). Transactions costs are a significant percentage of value of cattle in both types of markets (video and regional), but transactions costs are slightly lower in the video auction than in traditional auctions. This suggests that the video auction is somewhat more efficient in completing transactions than are traditional regional auctions, since cattle are usually shipped fewer miles for video auctions.

Adjusted video auction prices are higher than the OKC, Greeley, and Dodge City midpoint prices. While statistically significant, the difference in buyer prices between OKC and the video auction is small. Thus, video auction prices are basically equal to prices at what is widely recognized as the highest priced, most competitive feeder cattle regional auction. The video auction offers sellers the advantages of lower shrink than available when using traditional auctions and zero trucking costs.

Concluding Comments

Buyer and seller acceptance of video auctions is growing. Video auctions offer some unique features that should generate interest from industry, government, and the research community.

Satellite video auctions (as analyzed here) are national markets. Large numbers of cattle from numerous regional locations, with various weight and other characteristics are offered for sale. Consequently, these auctions could be a valuable source of market information.

Video auctions may provide buyers with more information about the history of cattle, type of feed, and vaccinations than traditional auctions

provide. One way of testing for relative adequacy of information within a market is to compare prices in different markets. This study found that net prices received by sellers and paid by buyers in the video auction exceeded reported regional market prices for the week following each video sale.

For many cattle producers, video auctions appear to provide higher net prices than regional auctions provide since trucking costs and shrink are reduced. Also, reduced health problems may allow buyers to offer higher prices for video auction cattle than for traditional auction cattle. Additionally, because video auctions reduce transportation costs, they represent a new source of competition with regional auctions where little within regional competition existed before. This is evidenced by the rapid growth of video auction sales volume. In fact, one concern is the impact of video cattle auctions on other markets, particularly on regional and local auctions. Video auctions are likely to shrink the market area of the regional auctions. Sellers close to the regional auctions will have lower trucking costs than those in distant locations and may still receive a higher net price in the regional market than through video auctions. Sellers with high trucking costs will find video auctions an attractive alternative to regional auctions.

[Received February 1989; final revision received May 1990.]

References

- Bailey, DeeVon, and B. Wade Brorsen. "Price Relationships Between Utah and Other Regional Markets for Fed Cattle." Utah State University Exp. Sta. Res. Rep. No. 107, Jan. 1986.
- Buccola, Steven T. "An Approach to the Analysis of Feeder Cattle Price Differentials." *Amer. J. Agr. Econ.* 62 (1980):574-80.
- Chou, Ya-lun. *Statistical Analysis*, 2nd ed. New York: Holt, Rinehart, and Winston, 1975.
- Faminow, Merle D., and Russel L. Gum. "Feeder Cattle Price Differentials in Arizona Auction Markets." *West. J. Agr. Econ.* 11(1986):156-63.
- Hamn, Shannon R., Wayne D. Purcell, and Michael A. Hudson. "A Framework for Analyzing the Impact of Anonymous Bidding on Prices and Competition in a Computerized Auction." *N. Cent. J. Agr. Econ.* 7 (1985):109-17.
- Janssen, Larry, and Richard Shane. "Price Variability at South Dakota Livestock Auction Markets." Econ. Dep. Res. Rep. No. 89-2, South Dakota State University. June 1989.
- Just, Richard E., and Gordon C. Rausser. "Commodity Price

- Forecasting with Large-Scale Econometric Models and the Futures Market." *Amer. J. Agr. Econ.* 63(1981): 197-208.
- Minish, Gary L., and Danny G. Fox. *Beef Production and Management*. Reston VA: Reston Publishing Co. 1979.
- Oellermann, Charles M., B. Wade Brorsen, and Paul L. Farris. "Price Discovery for Feeder Cattle." *J. Futures Mkts.* 9(1989):113-21.
- Pate, F. M. "Preconditioning Feeder Calves." Agr. Res. and Education Ctr., Florida Agr. Exp. Sta.
- Pollreis, J. P. "Arrival Procedures for Incoming Feedyard Cattle." *Animal Health and Nutrition*, Nov. 1986.
- Rhodus, W. Timothy, E. Dean Baldwin, and Dennis R. Henderson. "Pricing Accuracy and Efficiency in a Pilot Electronic Hog Market." *Amer. J. Agr. Econ.* 71 (1989):874-82.
- Scharlier, Marj. "Televised Auctioning of Livestock Could Spell the End for Stockyards." *Wall Street J.*, 2 Nov. 1988.
- Schroeder, Ted, James Mintert, Frank Brazle, and Orlen Grunewald. "Factors Affecting Feeder Cattle Price Differentials." *West. J. Agr. Econ.* 13(1988):71-81.
- Schultz, Robert W., and John M. Marsh. "Steer and Heifer Price Differences in the Live and Carcass Markets." *West. J. Agr. Econ.* 10(1985):77-92.
- Sporleder, Thomas L. "Implications of Electronic Trading for the Structure of Agricultural Markets." *Amer. J. Agr. Econ.* 66(1984):859-63.
- Sporleder, Thomas L., and Kathleen A. Mahoney. "Allocative Efficiency in Electronic Marketing for Feeder Cattle." Selected paper presented at AAEE annual meeting, Logan UT, 1-4 Aug. 1982.
- U.S. Department of Agriculture. Agricultural Marketing Service. Weekly reported price ranges from LS-214 forms for Greeley, OKC, and Dodge City for all weeks in 1987.
- Ward, Clement E. *An Analysis of Oklahoma Alfalfa Prices from Haymarket and Satellite Haymarket*. Oklahoma State University Agr. Exp. Sta. Professional Pap. No. PP 2898, 1989.

Evaluating Prior Beliefs in a Demand System: The Case of Meat Demand in Canada

James A. Chalfant, Richard S. Gray, and Kenneth J. White

An almost ideal demand system for meats is estimated using Canadian data. A Bayesian approach is used to impose inequality restrictions on substitution elasticities, via Monte Carlo integration and importance sampling, in order to conform with prior beliefs about curvature and monotonicity restrictions and substitution relationships. Results are more consistent with the concavity and monotonicity restrictions from demand theory than with the added restriction that all meats are substitutes.

Key words: almost ideal demand system, Bayesian inference, inequality constraints, meats demand.

The widespread use of flexible functional forms in demand analysis has given researchers the ability to model consumer preferences with no restrictions on the nature of substitution or complementarity relationships between pairs of goods. Unfortunately, theoretical restrictions automatically met by simpler forms need not hold with flexible forms, so estimated demand systems often conflict with prior beliefs. Symmetry and homogeneity restrictions may be violated when tested, but they are generally imposed using equality restrictions on the parameters of the model.

More difficult are restrictions represented as inequality constraints. It is common to observe predicted budget shares outside the 0–1 interval, violating monotonicity, or violation of curvature restrictions, reflected in a matrix of elasticities of substitution between goods that is not negative semidefinite. Thus, the researcher who desires to approximate arbitrary preferences often ends up with an approximation that suggests that they are badly behaved.

In this paper we show how to impose the inequality constraints of monotonicity and concavity of the consumer's expenditure function, using an almost ideal demand system for per capita consumption of meats and fish in Canada. While one might expect that many of the factors affecting U.S. consumption have also been present in Canada, relative prices have not always been similar. In particular, while beef and pork are, to a large extent, freely traded between the two countries, live poultry and poultry meat are not. During the early 1970s, Canada introduced a supply management scheme under Article 11(c) of GATT. Under this scheme, domestic prices have been supported above U.S. prices through production quotas for producers within Canada and import quotas to restrict trade. As a result, some of the relative decline in relative poultry prices that has occurred in the United States since 1970 was not observed in Canada.

In addition to the inequality restrictions from consumer theory, we suggest a new set of inequality restrictions not generally imposed on a demand system, but which have a compelling motivation. It seems reasonable to expect that no pair of foods that play essentially the same role in the diet should be complements. One thinks of coffee and cream or beef and gravy, but not beef and fish as complementary items. Yet, complementarity is common when flexible forms are estimated. In many applications, an intermediate case between the rigidity of Cobb-Douglas or CES preferences and the flexible form would be desirable.

James A. Chalfant is an associate professor of agricultural and resource economics, University of California, Berkeley; Richard S. Gray is an assistant professor of agricultural economics, University of Saskatchewan; and Kenneth J. White is a professor of economics, University of British Columbia.

The authors wish to thank many individuals, without implicating them, for helpful comments and discussions on this topic, especially the referees, Julian Alston, Tulay Bayri, Charles Blackorby, Ruey-er Chang, Erwin Diewert, John Geweke, William Griffiths, George Judge, Max King, Jeffrey Perloff, Dale Poirier, Linda Robbins, Dave Ryan, Gene Savin, Dan Sumner, Terry Wales, and Nancy Wallace.

While inequality restrictions on the signs of elasticities of substitution are suggested by empirical observation, rather than the underlying theory, they seem just as important as compatibility with theory in judging the degree to which an estimated demand system conforms to prior beliefs about consumer behavior. For applications such as this one, then, the constraint that all meats are substitutes can be viewed as another requirement that any well-behaved demand system must satisfy.

We make use of a Bayesian approach based on Geweke's (1986, 1988, 1989) work. He shows how to make inferences about or impose inequality restrictions in regression models. Chalfant and White used his method to impose curvature and monotonicity restrictions on the translog cost function but did not address restrictions on individual elasticities of substitution.

There are familiar alternatives to imposing prior beliefs with the Bayesian approach. One can always search over flexible forms to find one that is consistent with the desired restrictions but with the obvious consequences for inferences once a well-behaved demand system has been obtained (Judge and Bock). Alternatively, inequality restrictions can be imposed through constraints on a maximum-likelihood estimation procedure. Unfortunately, such constraints are difficult to interpret statistically (e.g., Gallant and Golub, Hazilla and Kopp, Wolak); the usual likelihood ratio test does not apply, for instance. Even if testing the restrictions is not the goal, imposing inequality constraints in this manner is likely to yield parameter values that lie on one of the constraints. For instance, constraining a demand elasticity to be nonpositive may well produce a vertical demand curve if the constraint is binding. Thus, a constrained maximum-likelihood approach is neither intuitively satisfying nor statistically attractive.

The next section discusses the almost ideal demand system. The Bayesian approach to inequality restrictions then is described in detail, followed by the application to meats demand. The data are described and demand system estimates are obtained for curvature and monotonicity restrictions alone and then with the added restriction that all meats are substitutes.

The Almost Ideal Demand System

Demand theory suggests that the demand for a good should be a function of its price, the prices

of other goods, and income. In order to estimate demand relationships in a system of a reasonable size, it is common to invoke weak separability; choices concerning the allocation of expenditures among a subset of goods consumed are assumed to be made independently from the prices of goods outside that group. For example, the quantity of beef consumed is likely to be a function of the prices of beef, pork, chicken, fish, and total expenditure on meat but is not a function of the price of bananas.¹

Whether or not it is appropriate to assume separability for a particular demand system is an empirical question. Theory suggests that any partial demand system representing the separable parts of larger systems should satisfy the conditions of symmetry, homogeneity, monotonicity, and concavity. Indeed, one interpretation of these conditions not holding in an estimated system is that the goods included in the demand system do not make up a separable group—a relevant price has presumably been omitted. Violation of those conditions also may indicate the presence of structural change, aggregation bias, or some other specification error.

Deaton and Muellbauer suggested the almost ideal demand system as a particular representation of price-independent, generalized logarithmic (PIGLOG) preferences. Such preferences are consistent with the aggregation of individual preferences. In addition, the functional form they chose is locally flexible in the sense used by Diewert and Barnett—it can attain arbitrary values for substitution elasticities at a given set of prices.

The equations for budget shares take the following form:

$$S_i = \alpha_i + \sum_{j=1}^M \gamma_{ij} \ln P_j + \beta_i \ln (x/P),$$

where P_j is the j th good's price, x denotes total expenditure on the M goods, and P is a price index. The specification of the underlying expenditure function introduced by Deaton and Muellbauer leads to the expression

$$\begin{aligned} \ln P = & \alpha_0 + \sum_{i=1}^M \alpha_i \ln P_i \\ & + 1/2 \sum_{i=1}^M \sum_{j=1}^M \gamma_{ij} \ln P_i \ln P_j, \end{aligned}$$

¹ Except, of course, to the extent that the prices of goods outside the group under study affect the total group expenditure, perhaps in a preliminary stage of allocating expenditures to aggregates such as meats, other foods, shelter, etc.

but this model, which is nonlinear in the parameters, is usually not estimated (for an exception, see Georgantelis, Phillips, and Zhang). More common is to follow the advice of Deaton and Muellbauer and replace this expression for $\ln P$ with Stone's geometric price index

$$\ln P = \sum_{k=1}^M S_k \log P_k$$

giving rise to an approximate almost ideal demand system that Blanciforti and Green (1983a, b) termed the linear approximate model because it is linear in parameters.

For either model, a system of these share equations can be estimated to obtain parameter estimates, and simple formulas convert the parameter estimates to elasticities. The system is easily restricted to satisfy symmetry ($\gamma_{ij} = \gamma_{ji} \forall i, j$) and homogeneity ($\sum_{j=1}^M \gamma_{ij} = 0 \forall i$). The adding-up property holds, given these restrictions, provided that $\sum_{i=1}^M \alpha_i = 1$ and $\sum_{i=1}^M \beta_i = 0$. Concavity or monotonicity restrictions are more difficult because they involve multiple inequality restrictions on the parameters. For concavity, the matrix of second derivatives of the expenditure function, or equivalently, of elasticities of substitution, must be negative semidefinite. For monotonicity, predicted budget shares must all be between 0 and 1 to ensure that predicted quantities consumed are positive.² Such restrictions are difficult to impose using most econometric packages and even harder to interpret statistically.

As with other flexible functional forms, elasticities are not constant with respect to prices and expenditures. The income elasticity of demand for good i is

$$\eta_i = 1 + \frac{\beta_i}{S_i}$$

for either specification of the price index. The Marshallian elasticity of demand for good i with respect to price j in the nonlinear model is

$$\begin{aligned} \sigma_{ij} &= -\delta_{ij} + \frac{1}{S_i} \cdot \left[\gamma_{ij} - \beta_i \alpha_j - \beta_i \sum_{k=1}^M \gamma_{jk} \ln P_k \right] \\ &= -\delta_{ij} + \frac{1}{S_i} \cdot \left[\gamma_{ij} - \beta_i \left(\alpha_j + \sum_{k=1}^M \gamma_{jk} \ln P_k \right) \right], \end{aligned}$$

where δ_{ij} equals 1 when $i = j$ and 0 otherwise.

The implied compensated price elasticity (η_{ij}^*) or elasticity of substitution (σ_{ij}) can then be obtained by manipulating the Slutsky equation in elasticity form:

$$\eta_{ij}^* = S_j \sigma_{ij} = \eta_{ij} + S_j \eta_i,$$

which yields, for the elasticity of substitution,

$$\begin{aligned} \sigma_{ij} &= 1 + \frac{\gamma_{ij}}{S_i S_j} + \frac{-\delta_{ij}}{S_i} \\ &\quad + \frac{\beta_i}{S_i} \left[1 - \frac{\alpha_j + \sum_{k=1}^M \gamma_{jk} \ln P_k}{S_j} \right]. \end{aligned}$$

The term in square brackets disappears if Stone's price index is used, leaving

$$\sigma_{ij} = 1 + \frac{\gamma_{ij}}{S_i S_j} + \frac{-\delta_{ij}}{S_i}.$$

Green and Alston noted that the latter formula is obtained for the linear-approximate model only if the budget shares appearing in Stone's price index are treated as exogenous, and they offered alternative formulas to accommodate this problem. However, the Monte Carlo results in Foster, Green, and Alston show that, if the nonlinear model is viewed as the underlying demand system and the linear-approximate model is indeed viewed as an approximation, the simpler formula provides a good approximation. This can be seen by noting that for homothetic models (where each $\beta_i = 0$), the two formulas are identical, and for nonhomothetic models, the discrepancy is likely small.

A final practical advantage of the simpler formulas is that they remain consistent with the income elasticity η_i above, whereas treating the budget shares in the price index as endogenous means that the model no longer will be characterized by these income elasticities. Deaton and Muellbauer noted that these elasticities were an advantage of the almost ideal model. In the application below, the simpler formulas are used, but the method would also apply with the "true" nonlinear model or the Green and Alston formulas.

The Bayesian Approach to Testing Inequality Restrictions

An alternative approach to imposing inequality restrictions in a demand system is made possible using a Bayesian approach, which permits the

² Chapter 3 of Varian, or a comparable text, provides further details concerning the restrictions from consumer theory.

formal inclusion of such prior information. Often, prior information can be imposed by choice of functional form. An extreme case is the Cobb-Douglas utility function, which would impose all of the restrictions from consumer theory plus some less desirable ones, such as additivity of preferences and that elasticities of substitution are each one. The latter restriction might lead one to question the stability of a demand system, for instance, if substitution patterns in response to relative price changes were not consistent with the assumed elasticities. Thus, flexible functional forms are a preferred alternative.

Symmetry and homogeneity restrictions from demand theory represent prior information that is often imposed on flexible forms through equality restrictions on the parameters. Such restrictions reduce the dimensionality of the parameter space when demand systems based on these forms are estimated; the symmetry and homogeneity restrictions, for instance, provide considerable gains in degrees of freedom. Prior information taking the form of an inequality restriction is less informative than such equality restrictions, in the sense that this information serves to truncate the parameter space, rather than reduce the number of free parameters. For instance, a particular parameter θ may be restricted to be positive. Conventional approaches to estimation do not permit the formal inclusion of such information (e.g., Judge et al.), and most econometric packages do not permit such restrictions to be imposed.

The problem of prior beliefs that take the form of inequality constraints is easily handled in the context of Bayesian inference. The Bayesian approach begins with a prior density function, defined over the vector of parameters, θ , call it $p(\theta)$.³ This prior density summarizes all of the information the researcher has about θ prior to estimation. Specifying $p(\theta)$ permits the formal inclusion of information about the parameters. For instance, if a particular parameter is considered equally likely to be positive or negative, a zero median characterizes the marginal probability density function used to describe prior beliefs about that parameter. If there is no prior information about θ , $p(\theta)$ is simply defined to be proportional to a constant over all real numbers, thus making it an improper density and a

so-called "diffuse prior." Alternatively, $p(\theta)$ could be a proper density function that reflects various beliefs about θ in the form of probability statements. A very simple case is the prior p.d.f. which says that θ is contained in some region D with probability one:

$$p(\theta) \propto c \quad \forall \theta \in D,$$

where \propto denotes "is proportional to"; D may be an open or closed interval, depending on the application. We consider below how such a prior density can be used to represent prior information about the parameters of a demand system.

Bayes' theorem shows how to combine prior and sample information to obtain a posterior distribution for the parameters in θ given a data set y :

$$f(\theta|y) \propto p(\theta) L(\theta|y),$$

where $L(\cdot)$ is the likelihood function based on the observed data. Unlike the sampling-theoretic approach to estimation, the Bayesian approach recognizes that posterior beliefs are conditional on the observed data set rather than emphasizing the performance of estimators in repeated samples.

The posterior distribution $f(\theta|y)$ summarizes all information available about θ , both prior and sample information. It can serve as the end result of an investigation, or it can be used to calculate confidence intervals and probabilities related to hypotheses about θ or to obtain a point estimate of θ or some related quantity such as a demand elasticity. The optimal point estimate for θ depends on the investigator's objective function. Constrained maximum-likelihood estimation, which yields the mode of the posterior distribution as a point estimate, corresponds to a "zero-one" loss function (e.g., Zellner 1988). This is an implausible loss function, in that it places the same weight on being wrong for estimates arbitrarily close to the true θ and for choices very far from that value. More reasonable loss functions can be imagined, and different point estimates will result. For instance, if the investigator's loss function is quadratic, the mean of the posterior distribution for θ minimizes expected loss (e.g., Judge et al.). All that is needed, then, to find Bayesian point estimates of the parameters of a demand system is a means to describe prior beliefs in the form of inequality restrictions using $p(\theta)$, a way to obtain the posterior density function, and then a way to find its mean.

Below we illustrate this approach using the quadratic loss function. With the inequality re-

³ See Zellner (1971) or Judge et al. for much more detailed descriptions of the properties of the Bayesian approach. An excellent introduction to inequality restrictions in the Bayesian framework is provided by Griffiths.

restrictions imposed, it is straightforward to obtain the mean of the posterior distribution, call it $\bar{\theta}$. This serves as the optimal point estimate of the parameters of the demand system.

Also of interest is some measure of the plausibility of the restrictions, given the data. Suppose prior beliefs are completely uninformative; that is, all parameter values are considered equally likely. In this case, the sample information dominates the posterior density function and an optimal point estimate is the mean of the (unrestricted) posterior distribution. The probability that the restrictions are correct can be calculated using the unconstrained posterior density. This probability is interpreted as the degree of belief that the restrictions are true, based on observed data, found by obtaining the probability that θ lies in D .

Both restrictions on the signs of substitution elasticities and restrictions on the entire matrix of substitution elasticities, to satisfy curvature restrictions, can be examined by calculating substitution elasticities. To evaluate these inequality restrictions, then, the behavior of elasticities of substitution must be examined everywhere in the parameter space where the researcher wishes to impose them. Similarly, the monotonicity restriction can be evaluated using predicted budget shares. Each set of restrictions is then imposed by truncating the parameter space so that each restriction holds. To obtain a Bayesian point estimate (with a quadratic loss function), the researcher must find the mean of the truncated posterior distribution for the parameter vector.

While these calculations are in principle straightforward, requiring that integrals over the posterior density function be evaluated, the analytic solutions cannot be obtained in practice, except for fairly simple models. The dimension of the posterior density is likely to be too great, even if the density function and the region of the parameter space of interest can be described easily. Instead, it is necessary to evaluate the integrals using Monte Carlo integration. This permits estimating the solutions to integrals by random sampling.⁴

To describe the method, we begin by specifying a data-generating process. We assume that prices and expenditures may be treated as exogenous, so that the parameters of the system

of $M - 1$ equations for budget shares could be estimated using seemingly unrelated regressions (SUR). As is well known, the equation for the M th budget share cannot be included without implying a singular contemporaneous covariance matrix for the error terms in the M share equations (Earten), but deleting the M th share and using restrictions on the parameters allows the complete set of parameter estimates to be obtained. Use of iterated SUR was shown by Barten to lead to maximum-likelihood estimates that are invariant to the equation chosen for deletion.

We assume that each time period's $M - 1$ vector of errors, and therefore the vector of budget shares, follows the multivariate normal distribution. Strictly speaking, one might prefer a distribution more compatible with the fact that observed shares are bounded by 0 and 1 (e.g., Woodland, Rossi), in which case imposing monotonicity is not an issue, but we prefer to stick with the more widely used distribution to illustrate the method. Our approach could easily be adjusted for nonnormal errors.

To illustrate Monte Carlo integration, suppose that Σ , the variance-covariance matrix of the errors, is known. Suppose also that $p(\theta)$, the prior information about the α_i 's, γ_{ij} 's, and β_i 's, indicates that some region D , a proper subset of R^J , contains the true parameter vector, where J denotes the number of free parameters in the model and R^J denotes the J -dimensional real numbers. Finally, suppose that the investigator has a quadratic loss function, and desires a point estimate of θ ; as noted earlier, the mean of the posterior density for θ minimizes expected loss.

The steps involved in finding an estimate of the mean, $\bar{\theta}$, are straightforward. With no prior information about θ and a known Σ , the posterior distribution for θ estimator would be the multivariate normal, centered at $\hat{\theta}$ with variance-covariance matrix $V(\hat{\theta})$, where $\hat{\theta}$ and $V(\hat{\theta})$ are obtained using Σ -restricted SUR. Given the prior information, the posterior distribution for θ then becomes the truncated multivariate normal, since θ is known to lie in D . The task becomes finding the mean of a truncated, J -variate normal:

$$E(\theta) = \int_{\theta_1 \in D} \dots \int_{\theta_J \in D} \theta N_J[\theta | \hat{\theta}, V(\hat{\theta})] d\theta_1 \dots d\theta_J.$$

Needless to say, such a calculation is infeasible for all but trivial examples.

Monte Carlo integration is based on the idea that an expectation such as the one above can

⁴ Kloek and van Dijk, van Dijk and Kloek, and Geweke (1986, 1988, 1989) provide the foundations for the Monte Carlo integration and importance sampling, described below. The application to demand systems follows the discussion in Chalfant and White. Gallant and Monahan; and Barnett, Geweke, and Yue have also followed a Bayesian approach to demand system estimation.

be estimated (arbitrarily accurately, given the Law of Large Numbers) using random sampling. One way of estimating the mean of a random variable with p.d.f. $f(\theta)$ is to generate a large number of replications in a random sample from that distribution, and calculate

$$\bar{\theta} = \frac{\sum_{i=1}^N \theta_i}{N},$$

where N is the number of replications. $\bar{\theta}$ serves as an estimate of $E(\theta)$, of course. Since N is determined by the investigator, $E(\theta)$ can be estimated with an arbitrarily high degree of accuracy. To apply that approach using the multivariate normal would require five steps:

(a) Estimate the parameters of the share equations, obtaining $\hat{\theta}$ and $V(\hat{\theta})$.

(b) Treat these as parameters of the posterior distribution for θ that would be consistent with no restrictions on θ 's range, the J -variate normal density.

(c) Use $N_J[\hat{\theta}, V(\hat{\theta})]$ and a random number generator to obtain a random sample from this multivariate normal. Omit those draws θ_i which are not contained in D , leaving a random sample of size n from the truncated multivariate normal.

(d) Estimate $E(\theta)$ using the average of the n replications in D :

$$\bar{\theta} = \frac{\sum_{i=1}^n \theta_i}{n}.$$

(e) A byproduct of the procedure is that $\hat{p}_D = n/N$ estimates the area under the multivariate normal density contained in D , i.e., the probability that the restrictions hold, given no prior information. If either p_D or $E(\theta)$ are estimated with less than the desired precision, increase N and repeat the process.

While somewhat computer-intensive, these steps are certainly feasible. They can be performed using the commonly available statistical packages for any desired posterior distribution. All that is required is a random number generator and some simple calculations.

Importance Sampling for Exact Results

The procedure outlined above relies on the asymptotic properties of the estimation procedure by making use of a normal approximation, since the covariance matrix Σ will invariably be unknown. While this is comparable to what is

done using non-Bayesian approaches, it will not yield results consistent with the exact posterior density function. Unfortunately, the posterior distribution for θ is no longer of the multivariate normal form when the variance matrix Σ is unknown. Effectively, the procedure outlined above substituted a conditional distribution $f(\theta|y, \hat{\Sigma})$ for the marginal distribution $f(\theta|y)$.

Such a marginal distribution for θ can be obtained from a joint posterior density for θ and the parameters of Σ by integrating over a posterior density for Σ . Following Zellner (1971, p. 242), Judge et al. (p. 478), or Srivastava and Giles, with a diffuse prior density for both Σ and θ , the resulting posterior density for θ is given by

$$f(\theta|y) \propto |A|^{-T/2},$$

where T denotes the number of sample observations. A typical element a_{ij} of the $M - 1$ by $M - 1$ matrix A is given by

$$a_{ij} = [(e_i(\theta))'(e_j(\theta))],$$

where $e_i(\theta)$ is the vector of residuals for share equation i evaluated using any value of θ where the posterior density is defined. This density function corresponds to the posterior density for θ with no prior information about likely values. Should one wish to impose the restriction that θ could take on only values consistent with the inequality restrictions, a truncated version of this posterior density must be used. In this framework, imposing the Bayesian restrictions and finding a posterior mean $\bar{\theta}$ requires sampling from the truncated density $f^R(\theta|y)$, but this is not a familiar density, so it is difficult to obtain an appropriate random sample. The procedure outlined above, whereby the untruncated density $f(\theta|y)$ could be used, also cannot be applied, for the same reason.

Instead, the steps outlined above must be modified, correcting for the fact that the multivariate normal is at best only approximately the correct posterior density. The technique for doing so is called importance sampling (e.g., Kloek and van Dijk, van Dijk and Kloek, Geweke 1986, 1988, 1989). The concept which underlies importance sampling is relatively straightforward. Before returning to the problem at hand, its use is illustrated with a simple example.

Consider estimating the mean of Z , a beta random variable with $\alpha = 9$ and $\beta = 1$ and density function $f(z)$. Such a random variable has

$$E(Z) = \int_0^1 z f(z) dz = \frac{\alpha}{\alpha + \beta} = .9.$$

Of course, this is an example where the integral could be evaluated and where random draws can also be obtained from the correct distribution, but suppose that we had available only a uniform random number generator, whose density function we denote $g(z)$. How might the mean of the beta distribution be calculated, using random draws from $U(0, 1)$?

The sample mean of replications drawn from the $U(0, 1)$ will underestimate $E(Z)$ because it will tend toward one-half. The reason, of course, is that values close to zero for Z occur more often under $U(0, 1)$ and most values above one-half will occur less often than under the beta distribution we have chosen. Importance sampling corrects for this by adjusting the "importance" given to each replication. The appropriate weight for each replication z_i is the ratio of the probability density function of the beta distribution at z_i , $f(z_i)$, and the density of the uniform at z_i , $g(z_i)$. In this way, those values drawn which are closer to one will receive a large weight while those closer to zero will receive a smaller weight.

To see why this works, note that the expected value of Z using the density function given by $f(z)$ can be found by integrating over $g(z)$ instead:

$$E(Z) = \int z f(z) dz = \int z \frac{f(z)}{g(z)} g(z) dz.$$

In the first instance, $E(Z)$ is taken with respect to f ; and in the second, $E[Z f(Z)/g(Z)]$ is taken with respect to g . Just as $E(Z)$ could be estimated using a sample mean of replications from $f(z)$, then so could it be estimated by drawing from $g(z)$ and calculating

$$\frac{\sum_{i=1}^N z_i \frac{f(z_i)}{g(z_i)}}{N}.$$

One surprising aspect of this procedure is that any density function can be used as $g(z)$, provided it is strictly positive over the range of Z determined from $f(Z)$. Otherwise, division by $g(z)$ within the integral is not allowed, and the implication would be that some values of Z , which do occur when sampling from $f(z)$, would never be drawn using $g(z)$. Naturally, if the weights applied to each z_i are close to one, so that $f(z)$ and $g(z)$ are similar, fewer draws will be required to obtain good estimates of the values of these integrals (Kloek and van Dijk; van Dijk and Kloek; Geweke 1986, 1988, 1989).

This procedure can be applied for estimating

the mean of the exact posterior density for θ . The posterior distribution is analogous to the beta distribution in the case above, in that it is the correct density but difficult to work with. Meanwhile, the multivariate t will be used as was the uniform distribution, to generate replications for θ . In the case of uninformative priors, draws from the multivariate t can be adjusted by the ratio of the two density functions to obtain an estimate of the probability that the inequality restrictions hold. Alternatively, the inequality restrictions can be imposed, so that the posterior density is truncated or restricted.

To find the posterior mean, find the solution to

$$E(\theta|y) = \int_{\forall \theta \in D} \theta f^R(\theta|y) d\theta.$$

At the same time, to calculate the probability that the restrictions hold, find

$$p = \int_{\forall \theta \in D} I(\theta) f(\theta|y) d\theta,$$

where $I(\theta) = 1$ if the restrictions hold and $I(\theta) = 0$ otherwise. Each of these could be accomplished by sampling from the exact posterior density $f(\theta|y)$, if it were of a known form, and the steps outlined earlier could be used. Because these integrands are too complicated to permit analytic solutions, importance sampling must be used. To reiterate, notice that the posterior mean can also be found by

$$E(\theta|y) = \int_{\forall \theta \in D} \theta \frac{f^R(\theta|y)}{g^R(\theta|y)} g^R(\theta|y) d\theta,$$

where $g^R(\theta|y)$ is the truncated multivariate t p.d.f. and D is the region of the parameter space consistent with the concavity and monotonicity restrictions. Again, the density $g^R(\theta|y)$ must be positive over the entire range (D) of the posterior density $f^R(\theta|y)$.

The modified steps now required for the calculations, taken from Chalfant and White, are given below:

(a) Estimate the parameters of the demand system using iterated seemingly unrelated regressions to obtain maximum likelihood estimates $\hat{\theta}$ and the estimated covariance matrix $V(\hat{\theta})$.

(b) Calculate ε matrix H such that $HH' = V(\hat{\theta})$. Draw a random vector of the same length as $\hat{\theta}$ from the standard normal distribution

$$w \sim N(0, I),$$

where I is the identity matrix of order J . Replications of θ that follow a multivariate t -distribution with λ degrees of freedom can be generated using the steps in van Dijk and Kloeck. We obtain

$$\theta^A = \hat{\theta} + Hw/v$$

and its "antithetic replication"

$$\theta^B = \hat{\theta} - Hw/v,$$

where v is obtained by taking a vector z of λ draws from $N(0, 1)$ and calculating

$$v = [(z'z)/\lambda]^{1/2}.$$

The inclusion of antithetic replications was suggested by Geweke (1988) to improve convergence.

Using the multivariate t ,

$$g(\theta|y) \propto \left[\lambda + (\theta - \hat{\theta})' \Sigma^{-1} (\theta - \hat{\theta}) \right]^{-(\lambda+J)/2}$$

(c) Check each replication to see if it violates concavity, monotonicity, or substitutability. To do so, we calculated elasticities of substitution using each replication and the means of observed budget shares for the four meats. Concavity and substitutability were checked using these elasticities. (To check these restrictions for all 29 observations would have involved simply repeating the same check for each observed share vector.) To check monotonicity, we used each replication to obtain new predicted shares for all twenty-nine data points.

(d) Estimate the mean of the posterior distribution using the n draws of θ^A or θ^B that satisfy the restrictions

$$\bar{\theta} = \frac{\sum_{k=1}^n \theta_k \frac{f^R(\theta_k|y)}{g^R(\theta_k|y)}}{\sum_{k=1}^n \frac{f^R(\theta_k|y)}{g^R(\theta_k|y)}}.$$

As noted by Chalfant and White, if $f(\theta|y)$ and $g(\theta|y)$ were proper density functions, a denominator of N would suffice. Otherwise, the denominator serves as a normalizing constant to correct for the fact that we use only the kernels of proper densities.

(e) To estimate the probability that the restrictions hold, use all replications, letting the first n be those consistent with the restrictions, and calculate

$$\hat{p}_D = \frac{\sum_{k=1}^N I(\theta_k) \frac{f(\theta_k|y)}{g(\theta_k|y)}}{\sum_{k=1}^N \frac{f(\theta_k|y)}{g(\theta_k|y)}} = \frac{\sum_{k=1}^n \frac{f^R(\theta_k|y)}{g^R(\theta_k|y)}}{\sum_{k=1}^N \frac{f(\theta_k|y)}{g(\theta_k|y)}}.$$

(f) Check to see if the number of replications is large enough to arrive at stable estimates of $\bar{\theta}$ or \hat{p} and of their standard deviations, if desired. If not, increase N .

A measure of the numerical accuracy of each estimated expectation, analogous to the usual standard error of the estimate of a population mean, was suggested by Geweke (1989). These can be used in step (f). The numerical standard error (nse) of the estimate of the mean of the posterior is given by

$$nse(\theta) = \frac{\sum_{k=1}^N (\theta_k - \bar{\theta}_k)^2 \left(\frac{f(\theta_k|y)}{g(\theta_k|y)} \right)^2}{\left(\sum_{k=1}^N \frac{f(\theta_k|y)}{g(\theta_k|y)} \right)^2}.$$

As also pointed out by Barnett, Geweke, and Yue, these measures are reliable estimates of the true nse's if and only if the tails of the distribution being sampled from are at least as "fat" as the tails of the posterior distribution. This characteristic of the importance sampling procedure is easily checked by plotting the replications of each θ_k against the ratio $f^R(\theta|y)/g^R(\theta|y)$. The weights should approach zero in the tails of each posterior distribution.

The standard deviation of the posterior distribution can be estimated by taking the square root of the estimated variance of the posterior distribution, obtained by using importance sampling to estimate $E[(\theta_k - \bar{\theta}_k)^2]$:

$$s.d.(\theta) = \left[\frac{\sum_{k=1}^N (\theta_k - \bar{\theta}_k)^2 \left(\frac{f^R(\theta_k|y)}{g^R(\theta_k|y)} \right)}{\sum_{k=1}^n \frac{f^R(\theta_k|y)}{g^R(\theta_k|y)}} \right]^{1/2}.$$

Application to Aggregate Meat Consumption in Canada

In this section, an almost ideal demand system for meat and fish products is estimated using aggregate Canadian data for the years 1960 to

1988.⁵ The data are from Agriculture Canada (Robbins) and were also used by Alston and Chalfant. The demands for four goods were examined—beef, pork, poultry, and fish (henceforth, the meats group). Consumer preferences for the meats group were assumed weakly separable from all other goods, and the demand system was treated as the second-stage allocation model, conditional on prices and a predetermined level of expenditures.⁶ Using the Agriculture Canada data on price indices and per capita quantities consumed for each meat, along with 1981 nominal prices reported in Van Kooten, we converted indices into current prices and calculated total expenditures on the group.

One remaining variable was included, a "trend" term taking the value 1 in 1975 and increasing by 1 in every subsequent period. Alston and Chalfant found that, although this data set is consistent with the hypothesis of stable preferences, the almost ideal demand system cannot be fit to these data without some allowance for dynamic influences; we chose the trend term as being more interpretable, albeit conditional on this specification, than an autocorrelation correction. A trend that began in the middle of the sample, 1975, seemed to correspond more to a hypothesized structural change or other type of model failure than did one beginning in 1960.

The share equations of the linear approximate form, obtained by deflating nominal expenditures by Stone's geometric price index, were then augmented by including this trend coefficient in each share equation. Thus the expression for the share of the budget allocated to the *i*th meat is

$$S_i = \alpha_i + \sum_{j=1}^4 \gamma_{ij} \ln(P_j) + \beta_i \ln(x/P) + \tau_i t.$$

A system of three such equations was estimated using the nonlinear regression (NL) procedure of version 6.1 of SHAZAM (White et al.). The fourth equation was deleted because of singularity of the variance matrix for all four equations, and parameters of that equation were obtained through the homogeneity and symmetry

Table 1. Unconstrained Results, Almost Ideal Demand System

Parameter ^a	Estimate	Standard Error
α_{beef}	-0.502	0.157
$\gamma_{\text{beef} \cdot \text{beef}}$	0.091	0.019
$\gamma_{\text{beef} \cdot \text{pork}}$	-0.004	0.012
$\gamma_{\text{beef} \cdot \text{poultry}}$	-0.009	0.013
$\gamma_{\text{beef} \cdot \text{fish}}$	-0.078	
β_{beef}	0.216	0.039
τ_{beef}	-0.0029	0.0007
α_{pork}	0.732	0.144
$\gamma_{\text{pork} \cdot \text{pork}}$	0.040	0.0145
$\gamma_{\text{pork} \cdot \text{poultry}}$	0.0162	0.017
$\gamma_{\text{pork} \cdot \text{fish}}$	-0.052	
β_{pork}	-0.010	0.037
τ_{pork}	-0.0020	0.0008
α_{poultry}	0.211	0.238
$\gamma_{\text{poultry} \cdot \text{poultry}}$	0.016	0.037
$\gamma_{\text{poultry} \cdot \text{fish}}$	-0.023	
β_{poultry}	-0.0006	0.063
τ_{poultry}	0.0032	0.0011
$\gamma_{\text{fish} \cdot \text{fish}}$	0.152	
β_{fish}	0.116	
τ_{fish}	0.0016	

^a Estimates of parameters without standard errors were obtained from the equality restrictions for symmetry, homogeneity, and adding-up.

restrictions. The adding-up property is preserved by letting the sum of the four trend coefficients equal zero. By iterating over both the parameters and the error variance-covariance matrix, the estimates obtained are invariant to the equation chosen for deletion (Barten).

Estimates of the parameters are given in table 1. The trend coefficients do appear to contribute significantly to the model, as the likelihood-ratio test statistic for the restriction that they are all zero is 35.19 (as compared with a critical value of $\chi^2_{3, \alpha=.05}$ of 11.345). They imply that, if the almost ideal model is the correct specification, there has been a negative trend affecting beef and pork and a positive one affecting chicken and fish.⁷

Elasticities of substitution are reported in table 2. Compensated price elasticities and expenditure elasticities are reported in table 3, part A, while uncompensated Marshallian elasticities

⁵ Other applications of this system to agricultural data include Blanciforti, Green, and King; Chalfant; Hayes, Wahl, and Williams; and Mcschini and Meilke. Hayes, Wahl, and Williams also consider the restriction that all meats are substitutes, in an examination of Japanese meats demand.

⁶ Although LaFrance notes that treating expenditures as predetermined is inconsistent with the assumption that quantities are measured with errors, this is standard practice. The method we use does not depend on this assumption, however, and could accommodate a joint explanation of the first- and second-stage allocation decisions.

⁷ Because the results from the nonparametric test for stable preferences were consistent with the existence of a stable demand system, the trends may be picking up misspecification of the almost ideal model. One way to decide whether one should instead disregard those results and conclude that significant trends imply structural change rather than misspecification is to analyze the implied behavior of elasticities in a variety of models. Our analysis of just one demand system cannot answer that question definitively, but the procedure followed below provides a means for such an analysis to incorporate prior beliefs.

Table 2. Elasticities of Substitution: Posterior Mean Values

Elasticity	Unconstrained	Concavity	Concavity and Substitutability
$\sigma_{\text{beef beef}}$	-1.13	-1.158 [0.003] ^a (0.188)	-1.207 [0.004] (0.181)
$\sigma_{\text{beef pork}}$	0.953	0.934 [0.003] (0.158)	0.902 [0.004] (0.156)
$\sigma_{\text{beef poultry}}$	0.853	0.797 [0.004] (0.225)	0.735 [0.006] (0.221)
$\sigma_{\text{beef fish}}$	-0.049	0.072 [0.005] (0.316)	0.252 [0.007] (0.272)
$\sigma_{\text{pork pork}}$	-2.384	-2.405 [0.005] (0.282)	-2.471 [0.007] (0.280)
$\sigma_{\text{pork poultry}}$	1.268	1.186 [0.008] (0.406)	1.087 [0.011] (0.404)
$\sigma_{\text{pork fish}}$	0.013	0.147 [0.009] (0.466)	0.375 [0.011] (0.423)
$\sigma_{\text{poultry poultry}}$	-3.641	-3.876 [0.020] (1.092)	-3.627 [0.023] (0.994)
$\sigma_{\text{poultry fish}}$	0.549	0.965 [0.019] (1.014)	0.95 [0.020] (0.835)
$\sigma_{\text{fish fish}}$	-0.458	-1.216 [0.020] (1.181)	-1.772 [0.019] (0.795)

^a The numerical standard errors of the estimated posterior means are reported in square brackets, and the corresponding standard deviations are in parentheses.

ties are given in table 4, part A. Recall that the latter make use of a constant level of meats expenditure, not total expenditure. These elasticities were calculated at the mean budget shares observed in the sample and were obtained without the inequality constraints.

The negative own-elasticities of substitution are as one would expect. The mostly positive elasticities of substitution indicate that the meats tend to be substitutes for one another at the mid-point of the sample. The unrestricted results that are contrary to prior belief are the negative elasticity of substitution, indicating complementarity, between fish and beef that is observed at the sample mean shares. Also contrary to expectations was the small positive elasticity between fish and pork and the fact that this elasticity also was often negative when the elasticities of substitution were evaluated at the individual observations.

Concavity holds at the mean and for the last eighteen observations in the sample; however,

the own-elasticity of substitution for fish is positive for the first eleven data points, violating a necessary condition for concavity. For all but six observations, at least one elasticity of substitution between two goods was negative for either beef and fish or beef and pork, sometimes both at once. Finally, monotonicity holds with these estimates at every point in the sample.⁸

Using the Bayesian procedure described earlier, the probability that the monotonicity, concavity, and substitution restrictions hold for this demand system can be estimated. We followed the procedure outlined earlier to obtain a sample size of 10,000 replications (5,000 plus the antithetic replications) from the multivariate *t*-distribution with $\lambda = 4$ degrees of freedom, again

⁸ A further indication that this model requires trend effects is that the results without the trend terms are completely inconsistent with the concavity or substitution restrictions. When the model was estimated with the trend coefficients set equal to zero and the estimated parameters were used to calculate elasticities, every observation violated both restrictions.

Table 3. Compensated Demand Elasticities

Quantity <i>i</i>	Price <i>j</i>			
	Beef	Pork	Poultry	Fish
Part A. Unconstrained Case				
Beef	-0.393	0.234	0.169	-0.010
Pork	0.331	-0.586	0.252	0.003
Poultry	0.297	0.312	-0.723	0.114
Fish	-0.017	0.003	0.109	-0.095
Expend. Elasticities	1.614	0.582	0.966	0.499
Part B. Concavity Imposed				
Beef	-0.403	0.230	0.158	0.015
Pork	0.325	-0.591	0.235	0.031
Poultry	0.277	0.291	-0.769	0.200
Fish	0.025	0.036	0.191	-0.253
Expend. Elasticities	1.575	0.537	0.832	0.745
Part C. Concavity and Substitutability Imposed				
Beef	-0.420	0.222	0.146	0.052
Pork	0.314	-0.607	0.216	0.078
Poultry	0.256	0.267	-0.720	0.197
Fish	0.088	0.092	0.189	-0.368
Expend. Elasticities	1.524	0.462	0.860	0.893

using SHAZAM. The in-sample results for $\hat{\theta}$ and $V(\hat{\theta})$ were used to specify the parameters of this distribution. Only four degrees of freedom were chosen to ensure that sampling would occur from a density with relatively "fat" tails, thereby assuring that the entire range of the actual posterior distribution is covered by the importance

Table 4. Marshallian Demand Elasticities

Quantity <i>i</i>	Price <i>j</i>			
	Beef	Pork	Poultry	Fish
Part A. Unconstrained Case				
Beef	-0.955	-0.163	-0.151	-0.346
Pork	0.129	-0.729	0.136	-0.118
Poultry	-0.039	0.074	-0.914	-0.087
Fish	-0.191	-0.119	0.010	-0.199
Part B. Concavity Imposed				
Beef	-0.951	-0.158	-0.154	-0.312
Pork	0.138	-0.723	0.129	-0.081
Poultry	-0.012	0.087	-0.934	0.028
Fish	-0.234	-0.147	0.044	-0.407
Part C. Concavity and Substitutability Imposed				
Beef	-0.950	-0.153	-0.157	-0.264
Pork	0.153	-0.721	0.124	-0.018
Poultry	-0.044	0.056	-0.891	0.0187
Fish	-0.223	-0.127	0.011	-0.554

density. Monotonicity was checked by using each replication to predict budget shares for all twenty-nine observations; every replication was consistent with positive budget shares for each observation, so monotonicity was never violated. Concavity was checked for each replication by calculating substitution elasticities using the parameter values given by each replication and the mean observed budget shares. The eigenvalues were then calculated for each matrix of elasticities of substitution.

Consistency with the concavity restriction (a substitution matrix without positive eigenvalues) was found in over half of the replications. Had we been sampling from the exact posterior, the proportion consistent with concavity would have been the estimate of the probability that concavity holds, given diffuse priors. As noted earlier, however, that probability can be estimated using importance sampling to correct for not sampling from the exact posterior. Calculation of the ratio \hat{p}_D in step (e) of section 4 yielded a value of 0.57. These results imply that the demand system is reasonably well behaved by this criterion. Table 5 summarizes all of the results in this section for the posterior probabilities that various restrictions are true.

Elasticities of substitution calculated using the concavity restriction are given in table 2, with numerical standard errors and standard deviations. The point estimates for the elasticities are obtained by estimating the means of the posterior distributions for each elasticity. Those distributions were found using the posterior distributions for the parameters, the mean budget shares, and the formulas for elasticities. The own

Table 5. Posterior Probabilities of Restrictions

Inequality Restriction	None	Concavity	Concavity and Substitutability
Concavity (n.s.e.)	0.57 0.010		
All substitutes (n.s.e.)	0.6 0.003	0.28	
$\sigma_{\text{beef pork}} > 0$	1	1	
$\sigma_{\text{beef poultry}} > 0$	0.999	1	
$\sigma_{\text{beef fish}} > 0$	0.427	0.59	
$\sigma_{\text{pork poultry}} > 0$	0.599	1	
$\sigma_{\text{pork fish}} > 0$	0.507	0.63	
$\sigma_{\text{poultry fish}} > 0$	0.713	0.89	
$\tau_{\text{beef}} < 0$	1	1	1
$\tau_{\text{pork}} < 0$	0.593	1	1
$\tau_{\text{poultry}} > 0$	0.583	0.976	0.984
$\tau_{\text{fish}} > 0$	0.872	1	1
all τ	0.853	0.974	0.984

elasticities of substitution tend to be more negative than before (reflecting the concavity restriction), while imposing concavity alone was enough to make the average over all replications of each partial elasticity of substitution turn out to be positive. Expenditure and compensated price elasticities calculated using $\bar{\theta}$ are shown in table 3, part B, while table 4, part B contains corresponding uncompensated elasticities.

The system was less consistent with the restriction that all meats are substitutes. Imposing concavity and substitutability jointly reduces the parameter space D further; the probability that the joint restriction holds, given diffuse prior beliefs, falls to 0.16. (Alternatively, this can be viewed as a conditional probability of substitutability, given that concavity is imposed, of $0.16/0.57 = 0.28$.) This relatively stronger rejection of prior beliefs casts some doubt on the estimated system being a valid representation of preferences. Nonetheless, posterior means were calculated for the remaining elasticities. These numbers are shown in table 2 for the substitution elasticities and in part C of tables 3 and 4 for the price elasticities. For comparison to the results in table 1, the constrained parameter estimates are reported in table 6.

It is important to note that violations of the substitution restriction are not due solely to the cases of complementarity relationships implied by $\hat{\theta}$ —beef and fish or pork and fish. Any elasticity could be responsible. Unless the posterior density for any particular elasticity of substitution implies that it is positive with probability one, each can be responsible for any particular replication violating the constraint. We examined the posterior distributions for each elasticity and found that those two elasticities indeed were responsible for the lion's share of rejections. We never observed a replication that implied that beef and pork were complements at the mean budget shares, while the estimated probabilities that the substitution elasticity was positive for the beef-poultry and pork-poultry combinations were both greater than 0.99. In contrast, the probability that the elasticity of substitution was positive for the remaining cases of $\sigma_{\text{beef} \cdot \text{fish}}$, $\sigma_{\text{pork} \cdot \text{fish}}$, and $\sigma_{\text{poultry} \cdot \text{fish}}$ were 0.59, 0.63, and 0.89, respectively, suggesting that fish is most responsible for the violations.

One can also reexamine the trend results with an inequality-constrained model; although the trend coefficients do not affect elasticities directly, their point estimates are correlated with those of the γ_{ij} parameters and will be affected by imposing the inequality restrictions. The

Table 6. Inequality Constrained Coefficients

	Concavity	Concavity and Substitutability
α_{beef}	-0.4429 [0.0043] ^a (0.1766)	-0.3769 [0.0066] (0.1689)
$\gamma_{\text{beef beef}}$	0.0867 [0.0005] (0.0219)	0.0808 [0.0008] (0.0195)
$\gamma_{\text{beef pork}}$	-0.0056 [0.0003] (0.0134)	-0.0084 [0.0005] (0.0123)
$\gamma_{\text{beef poultry}}$	-0.014 [0.0004] (0.0152)	-0.0183 [0.0007] (0.0153)
$\gamma_{\text{beef fish}}$	-0.0671 0.2002 [0.0011] (0.0433)	-0.0541 0.1823 [0.0016] (0.0408)
β_{beef}		
τ_{beef}	-0.0032 [0.0000] (0.0008)	-0.0036 [0.0000] (0.0007)
α_{pork}	0.7856 [0.0044] (0.1569)	0.8555 [0.0049] (0.1252)
$\gamma_{\text{pork pork}}$	0.0401 [0.0004] (0.0169)	0.0361 [0.0007] (0.0155)
$\gamma_{\text{pork poultry}}$	0.0091 [0.0006] (0.0197)	0.0042 [0.0008] (0.0184)
$\gamma_{\text{pork fish}}$	-0.8348 -0.1138 [0.0011] (0.0400)	-0.8958 -0.1322 [0.0012] (0.0312)
β_{pork}		
τ_{pork}	-0.0022 [0.0000] (0.0008)	-0.0026 [0.0000] (0.0006)
α_{poultry}	0.3344 [0.0064] (0.2432)	0.3125 [0.0099] (0.2109)
$\gamma_{\text{poultry poultry}}$	0.0064 [0.0009] (0.0391)	0.0162 [0.0014] (0.0324)
$\gamma_{\text{poultry fish}}$	-0.3385 -0.0334 [0.0016] (0.0632)	-0.3261 -0.0279 [0.0026] (0.0547)
β_{poultry}		
τ_{poultry}	0.0025 [0.0000] (0.0012)	0.0025 [0.0000] (0.0010)
α_{fish}	0.3229	0.2089
$\gamma_{\text{fish fish}}$	1.2404	1.276
β_{fish}	-0.053	-0.0222
τ_{fish}	0.0029	0.0037

^a The numerical standard errors for each element of $\bar{\theta}$ are reported in square brackets, and the corresponding standard deviations are in parentheses. Parameters without these measures were obtained from the equality restrictions, as in table 1.

negative point estimates for beef and pork that were reported in table 1 for the trend effects observed without the inequality restrictions are less than the resulting posterior mean values for those two coefficients. The trend coefficients in both the beef and pork equations rise in absolute value, as does the implied one in the omitted fish equation, when concavity and then substitutability are imposed. At the same time, the coefficient in the poultry equation falls slightly. All of the estimated means of the posterior distributions for the trend coefficients are more than twice the corresponding standard deviations, as shown in table 6, although this is barely true for poultry. The posterior probability of a negative value in the beef and pork equations was 1, which was the probability of a positive value in the fish equation; only for poultry does the posterior distribution have some probability on both sides of zero, although that value is small for negative values, as the estimated probability of a positive value in the poultry equation is over .95.

A sample size of 10,000 was sufficiently large to obtain accurate estimates for these expectations from the posterior distribution. This is reflected in the small numerical standard error estimates, obtained as described in step (f). These are well within tolerable ranges for applications.

Summary and Conclusions

An unfortunate byproduct of the use of demand systems that do not restrict substitution elasticities is that theoretical restrictions such as symmetry or homogeneity are often violated. More difficult to cope with are inequality restrictions. The familiar problem of curvature or monotonicity restrictions is the best example, but the signs of elasticities of substitution between goods are also good examples. In order to determine whether an estimated demand system is entirely consistent with prior beliefs, it is important to be able to impose or make inferences about inequality restrictions.

We showed that a Bayesian procedure handled this problem nicely. It produces constrained parameter estimates and also an estimate of the probability that the restrictions are true. For the demands for beef, pork, chicken, and fish in Canada, some support was found for the concavity of the consumer's expenditure function underlying an almost ideal demand system ($p = 0.57$), while the monotonicity restric-

tion is definitely consistent with the model ($p = 1$). On the other hand, the sample information is less consistent with prior beliefs about substitution relationships; it reveals a low conditional probability (given concavity) that these four goods are all substitutes for each other ($p = .28$). Finally, whether or not these restrictions are imposed, the almost ideal demand system implies a negative trend affecting beef and pork consumption and a positive trend for chicken and fish.

Because the necessary integrals over the exact posterior density were quite complicated, Monte Carlo integration was used to estimate parameter values. The p.d.f. of the exact posterior was known but not recognizable, which made sampling from it difficult. This problem was overcome by importance sampling.

These findings concerning Canadian meat demand are conditional on the observed data and on the specification of the almost ideal demand system. All such inferences in demand systems are also conditional on separability and aggregation assumptions (Chalfant and Alston); but, if these assumptions are valid, the results must be interpreted as questioning either the prior belief that these goods are substitutes or the functional form for the demand system. If one is satisfied with the almost ideal model and unconcerned about the substitution restriction, the reasonably good results for concavity permit the use of the constrained elasticity estimates for policy analysis. The ability to produce a theoretically consistent set of elasticities is one of the main advantages of this procedure.

On the other hand, one might hesitate to use the estimated trends as evidence concerning taste changes if the lower probability for the restriction that all meats are substitutes is a cause for alarm. However, the results suggest that the almost ideal model cannot be estimated without the trend effects. Further research with other functional forms can help to address this question. It will be of interest to see how the probabilities we report change when another functional form is tried; Alston and Chalfant and the Monte Carlo studies to date certainly offer evidence that point estimates can change. In that light, the procedure and restrictions suggested in this study serve not only as a means to interpret the data but as a way of evaluating alternative functional forms for demand systems.

[Received January 1989; final revision received July 1990.]

References

- Alston, Julian M., and James A. Chalfant. "Can We Take the Con Out of Meat Demand Studies?" Dep. Agr. Econ., University of California, Davis, 1990.
- Barnett, W. A. "New Indices of the Money Supply and the Flexible Laurent Demand System." *J. Bus. and Econ. Statist.* 1(1983):7-23.
- Barnett, W. A., J. Geweke, and P. Yue. "Semiparametric Bayesian Estimation of the Asymptotically Ideal Model: the AIM Demand System." *Nonparametric and Semiparametric Methods in Econometrics*, ed. W. A. Barnett, J. Powell, and G. Tauchen. Cambridge: Cambridge University Press, forthcoming.
- Barten, Anton P. "Maximum Likelihood Estimation of a Complete System of Demand Equations." *Eur. Econ. Rev.* 1(1969):7-73.
- Blanciforti, L., and R. D. Green. "An Almost Ideal Demand System Incorporating Habits: An Analysis of Expenditures on Food and Aggregate Commodity Groups." *Rev. Econ. and Statist.* 65(1983a):511-15.
- . "The Almost Ideal Demand System: A Comparison and Application to Food Groups." *Agr. Econ. Res.* 35(1983b):1-10.
- Blanciforti, L., R. D. Green, and G. A. King. *U.S. Consumer Behavior over the Postwar Period: An Almost Ideal Demand System Analysis*. Giannini Foundation Monograph No. 40, University of California, Berkeley, 1986.
- Chalfant, James A. "A Globally Flexible, Almost Ideal Demand System." *J. Bus. and Econ. Statist.* 5(1987):233-42.
- Chalfant, James A., and Julian M. Alston. "Accounting for Changes in Tastes." *J. Polit. Econ.* 96(1988):390-410.
- Chalfant, James A., and Kenneth J. White. "Estimation of Demand Systems with Concavity and Monotonicity Constraints." Dep. Agr. and Resour. Econ., University of California, Berkeley, 1988.
- Deaton, Angus, and John Muellbauer. "An Almost Ideal Demand System." *Amer. Econ. Rev.* 70(1980):312-26.
- Diewert, W. E. "Applications of Duality Theory." *Frontiers of Quantitative Economics, Vol. II*, ed. M. D. Intriligator and D. A. Kendrick. Amsterdam: North-Holland Publishing Co., 1974.
- Foster, Kenneth, Richard D. Green, and Julian M. Alston. "Estimating Elasticities with the Linear Approximate Almost Ideal Demand System." Dep. Agr. Econ., University of California, Davis, 1990.
- Gallant, A. R., and G. H. Golub. "Imposing Curvature Restrictions on Flexible Functional Forms." *J. Econometrics* 26(1984):295-321.
- Gallant, A. R., and J. F. Monahan. "Explicitly Infinite Dimensional Bayesian Analysis of Production Technologies." *J. Econometrics* 30(1985):171-201.
- Georgantelis, S., G. D. A. Phillips, and W. Zhang. "Estimating and Testing an Almost Ideal Demand System." *The Practice of Econometrics*, ed. R. D. H. Heijmans and H. Neudecker. Dordrecht, Netherlands: Martinus Nijhoff Publishers, 1987.
- Geweke, J. "Antithetic Acceleration of Monte Carlo Integration in Bayesian Inference." *J. Econometrics* 38(1988):73-89.
- . "Bayesian Inference in Econometric Models Using Monte Carlo Integration." *Econometrica* 57(1989):1317-39.
- . "Exact Inference in the Inequality Constrained Normal Linear Regression Model." *J. Appl. Econometrics* 1(1986):127-41.
- Green, R. D., and J. M. Alston. "Elasticities in AIDS Models." *Amer. J. Agr. Econ.* 72(1990):442-45.
- Griffiths, W. E. "Bayesian Econometrics and How to Get Rid of Those Wrong Signs." *Rev. Mktg. and Agr. Econ.* 56(1988):36-56.
- Hayes, Dermot J., Thomas I. Wahl, and Gary W. Williams. "Testing Restrictions on a Model of Japanese Meat Demand." Paper presented at AAEA annual meeting, Knoxville TN, 31 July-3 Aug. 1988.
- Hazilla, Michael, and Raymond J. Kopp. "Testing for Separable Functional Structure Using Temporary Equilibrium Models." *J. Econometrics* 33(1986):119-41.
- Judge, G. G., and M. E. Bock. *The Statistical Implications of Pre-Test and Stein-Rule Estimators in Econometrics*. Amsterdam: North-Holland Publishing Co., 1978.
- Judge, George. G., W. E. Griffiths, R. Carter Hill, Helmut Lutkepohl, and Tsoung-Chao Lee. *The Theory and Practice of Econometrics*, 2nd. ed. New York: John Wiley & Sons, 1985.
- Kloek, T., and H. K. van Dijk. "Bayesian Estimates of Equation System Parameters: An Application of Integration by Monte Carlo." *Econometrica* 46(1978):1-19.
- LaFrance, J. "When Is Expenditure 'Exogenous' Separable Demand Models?" Dep. Econ., Montana State University, 1990.
- Moschini, G., and K. Meilke. "Modeling the Pattern of Structural Change in U.S. Meat Demand." *Amer. J. Agr. Econ.* 71(1989):253-61.
- Robbins, Linda. *Handbook of Prices, Consumption and Expenditures*. Ottawa: Agriculture Canada, 1989.
- Rossi, P. "Comparison of Alternative Functional Forms in Production." *J. Econometrics* 30(1985):345-61.
- Srivastava, V. K., and D. E. A. Giles. *Seemingly Unrelated Regression Models: Estimation and Inference*. New York: Marcel Dekker, 1987.
- Stone, J. R. N. *The Measurement of Consumers' Expenditure and Behavior in the United Kingdom, 1920-1938, Vol. I*. Cambridge: Cambridge University Press, 1954.
- Van Dijk, H. K., and T. Kloek. "Further Experience in Bayesian Analysis Using Monte Carlo Integration." *J. Econometrics* 14(1990):307-28.
- Van Kooten, G. C. "The Economic Impacts on Consumers of Government Intervention in the Poultry and Egg Sectors: A Comparison of Alternative Welfare Measures." Ottawa: Agriculture Canada Policy Branch Work. Pap. No. 5/87, 1987.
- Varian, Hal R. *Microeconomic Theory*. New York: W. W. Norton & Co., 1978.
- White, Kenneth J., S. Donna Wong, Diana Whistler, and Shirley A. Haun. *SHAZAM Econometrics Computer Program: User's Reference Manual, Version 6.2*. New

- York: McGraw-Hill Book Co., 1990.
- Wolak, Frank A. "Testing Inequality Constraints in Linear Econometric Models." *J. Econometrics* 41(1989):205-35.
- Woodland, A. D. "Stochastic Specification and the Estimation of Share Equations." *J. Econometrics* 10(1980):361-83.
- Zellner, A. "An Efficient Method of Estimating Seemingly Unrelated Regression and Tests for Aggregation Bias." *J. Amer. Statist. Assoc.* 57(1962):348-68.
- . *An Introduction to Bayesian Inference in Econometrics*. New York: John Wiley & Sons, 1971.
- . "Bayesian Analysis in Econometrics." *J. Econometrics* 37(1988):27-50.

Ex Post Flexibility and Choice of Capacity for Loss Reduction

L. Dean Hiebert

This paper examines the implications of *ex post* utilization flexibility for the firm's *ex ante* choice of capacity to supply a loss-reducing input. It is shown that the optimal *ex ante* choice is greater in the presence of flexibility than in the absence of flexibility. It is also shown that the total effect of an increase in risk on the *ex ante* choice includes a flexibility effect when *ex post* flexibility is present. The flexibility effect always tends to increase the demand for capacity.

Key words: damage control, *ex post* flexibility, loss reduction.

The firm's use of loss-reducing inputs has been widely studied. Feder analyzes the use of pesticides to reduce crop losses. Centner and Wetzstein consider the use of disease control measures by the buyers of livestock. Hiebert (1983) examines a firm's demand for self-insurance, an input which reduces the severity of the loss resulting from randomly occurring accidents. Lichtenberg and Zilberman discuss the specification of the production function when loss-reducing inputs are used.

This previous literature assumes that all loss containment decisions must be made before learning the actual severity of the realized state of nature. In certain decision-making contexts, however, some loss reduction decisions can be postponed until the realized state of nature is known to the decision maker. Suppose a firm must select a capacity to supply a loss-reducing input before learning the actual state of nature but can select the utilization of this capacity after learning the actual state of nature. In this case capacity is an *ex ante* decision variable, while the utilization of this capacity is an *ex post* decision variable.¹ For example, wind machines and sprinklers in citrus groves represent a capacity to supply frost protection services, but the

utilization of this capacity is chosen after the actual state of nature can be accurately predicted.

This paper considers the effect of *ex post* utilization flexibility on a firm's *ex ante* commitment to loss reduction. The firm makes an *ex ante* commitment by choosing a capacity to supply the loss-reducing input. If *ex post* utilization flexibility is present, the firm postpones the input choice until after learning the realized state of nature. In the absence of *ex post* utilization flexibility, the use of the input is fixed and equal to the capacity level. It is shown that the introduction of *ex post* flexibility always increases a firm's *ex ante* commitment to loss reduction. Moreover, the introduction of *ex post* flexibility can reverse the direction of the firm's *ex ante* capacity response to an increase in uncertainty.

Capacity Choice with Ex Post Flexibility

This section develops a simple model of capacity choice when *ex post* utilization flexibility is present. The firm selects a capacity level, k , before learning the actual realization of the state of nature. After observing the realized state of nature, the firm chooses the quantity of loss reduction services, x , given the capacity constraint $x \leq k$.

If a disaster occurs, the dollar loss is given by the damage function

$$(1) \quad D(x, \theta) \quad \text{for } x \leq k,$$

where θ is a random variable which represents the severity of the state of nature. The loss is an increasing function of θ . For all values of θ ,

L. Dean Hiebert is an associate professor of economics, Illinois State University.

The author thanks two anonymous reviewers for helpful comments.

¹ Turnovsky, Hartman, and Epstein discuss the *ex ante* and *ex post* decisions of a competitive firm facing product price uncertainty. Robison and Barry (chap. 17) consider a situation in which a "flexibility" characteristic of a durable is chosen before the selling price of the durable's services is known, while the quantity of services to be extracted can be chosen after learning the price.

the loss is a decreasing strictly convex function of x ($D_x < 0$, $D_{xx} > 0$), implying that loss reduction measures are subject to diminishing returns. The marginal effectiveness of the input will depend, in general, on the realized state of nature. Two alternatives are considered. In one case, the marginal effectiveness of the input is greater in more adverse states of nature ($-D_{x\theta} > 0$).² In the other case, the marginal effectiveness of the input is greater in more benign states of nature ($-D_{x\theta} < 0$).

Consider first the firm's *ex post* decision. Let $w(x)$ represent the variable cost of utilizing the installed capacity, where variable cost is an increasing convex function of x . The firm will choose x so that the marginal return to the input is greater than or equal to the marginal cost of the input. Thus, the optimal level of the input, x^* , must satisfy

$$(2) \quad -D_x(x^*, \theta) \geq w'(x^*) \quad \text{for } x^* \leq k.$$

Suppose that loss reduction measures are more effective in more severe states of nature. Then the optimal value of x is an increasing function of θ . If the state of nature is sufficiently severe, the strict inequality will hold in (2) and capacity is fully utilized, $x^* = k$. In less severe states of nature, the equality will hold in (2), and $x^* \leq k$. Hence, the optimal value of x can be written as

$$(3) \quad x^* = \min [x^*(\theta, w), k],$$

where $x^*(\theta, w)$ is the value of x which satisfies $-D_x(x^*, \theta) = w'(x^*)$.

At the time the firm makes the capacity decision, θ and hence x^* are unknown. The firm is risk averse and chooses a level of capacity to maximize the expected utility of *ex post* profit.³ Expected utility in any period can be written as

$$(4) \quad EU = (1-p)U(W-rk) + p \int_0^{\theta^*(k)} U(W-D(x^*, \theta)) - w(x^*) - rk dF$$

² In this case x is a risk-reducing input (as defined by Just and Pope) because the variance of profit is a decreasing function of x . Alternatively, if $-D_{x\theta} < 0$, the input is a risk-increasing input. In Hartman and Epstein, the inputs are risk-increasing because product price enters the profit function multiplicatively.

³ If the profit endowment and the price variables are identical in all periods and if θ is identically and independently distributed in all periods, then maximizing the present value of expected utility is identical to maximizing equation (4).

$$- p \int_{\theta^*(k)}^{\infty} U(W-D(k, \theta)) - w(k) - rk dF,$$

where U is a strictly concave utility function, W denotes the firm's profit endowment which is subject to the risk of partial loss D , r is the fixed cost of a unit of capacity, p is the probability that a loss occurs, and F is the distribution function of θ . Both W and r are assumed to be non-random constants; $\theta^*(k)$ denotes the critical value of θ for which the full utilization of capacity is the unconstrained *ex post* optimum. It is defined by $-D_x(k, \theta^*) = w$ and is a continuous function of the level of capacity.

The optimal level of capacity must satisfy the condition⁴

$$(5) \quad \partial EU / \partial k = (1-p)U'(\pi_0) \cdot -r - p \int_0^{\infty} U'(\pi_1) \cdot -rdF + p \int_{\theta^*(k)}^{\infty} U'(\pi_1)(-D_x(k, \theta) - w'(k))dF = 0,$$

where π_0 denotes profit if no loss occurs and π_1 denotes (random) profit if a loss actually occurs. This condition indicates that capacity is expanded until the marginal benefit (in expected utility) of increased capacity equals the marginal cost of capacity. The expectation of the marginal benefit is over those states of nature in which capacity is fully utilized because the benefit of added capacity is zero otherwise.

Alternatively, consider the case in which loss reduction is more effective in more benign states of nature. In this case the optimal input level is a decreasing function of θ . If the state of nature is sufficiently benign, the strict inequality will hold in (2) and $x^* = k$. In more severe states of nature the equality will hold and $x^* \leq k$. The firm's expected utility can be written as

$$(6) \quad EU = (1-p)U(W-rk) + p \int_0^{\theta^*(k)} U(W-D(k, \theta)) - w(k) - rk dF + p \int_{\theta^*(k)}^{\infty} U(W-D(x^*, \theta) - w(x^*) - rk) dF.$$

⁴ Condition (5) is also a sufficient condition for an optimum because EU is a strictly concave function of k .

The optimal level of capacity now satisfies

$$(7) \quad \partial EU / \partial k = (1 - p)U'(\pi_0) \cdot -r \\ + p \int_0^\infty U'(\pi_1) \cdot -rdF \\ + p \int_0^{\theta^*(k)} U'(\pi_1)(-D_x(k, \theta) \\ - w'(k))dF = 0.$$

Capacity Choice with Fixed Utilization

If the firm's use of x cannot be varied *ex post*,⁵ then $x = k$, and expected profit becomes

$$(8) \quad EU = (1 - p)U(W - rk) + p \int_0^\infty \\ \cdot U(W - D(k, \theta) - w(k) - rk)dF.$$

The optimal level of capacity must satisfy the condition

$$(9) \quad \partial E\pi / \partial k = (1 - p)U'(\pi_0) \cdot -r \\ + p \int_0^\infty U'(\pi_1) \cdot -rdF \\ + p \int_0^\infty U'(\pi_1)(-D_x(k, \theta) \\ - w'(k))dF = 0.$$

Again, the marginal benefit of added capacity is equated to its marginal cost. The expectation of the marginal benefit is over all states of nature because the firm is unable to vary the utilization of capacity *ex post*. In this case the *ex post* marginal net return to the loss reducing input, $-D_x(k, \theta) - w'(k)$, will be negative in some states of nature. This immediately implies that the optimal value of expected utility is greater in the presence of flexibility than in the absence of flexibility.

The Impact of Flexibility

This section compares the optimal choice when flexibility is present with that when flexibility is absent.

Suppose that loss reduction measures are more effective in more severe states of nature. Subtracting the left-hand side of condition (9) from the left-hand side of condition (5) and evaluating the result at the optimal choice under zero flexibility yields

$$(10) \quad -p \int_0^{\theta^*(k)} U'(\pi_1)(-D_x(k, \theta) - w'(k))dF > 0$$

because $-D_{x\theta} > 0$. This result implies that the firm selects a larger capacity in the presence of flexibility than in the absence of flexibility. It is because flexibility increases the marginal benefit of additional capacity.⁶

A similar result is obtained when loss reduction is more effective in less severe states of nature. Subtracting the left-hand side of (9) from the left-hand side of (7) yields

$$(11) \quad -p \int_{\theta^*(k)}^\infty U'(\pi_1)(-D_x(k, \theta) - w'(k))dF > 0$$

because $-D_{x\theta} < 0$. Again, the presence of flexibility increases the firm's demand for loss-reducing capacity. Hence, the impact of *ex post* flexibility on the optimal capacity choice does not depend on the risk-affecting character of the technology.

The Impact of Increased Uncertainty

This section considers the impact on the optimal capacity choice of increases in uncertainty as defined by Rothschild and Stiglitz.⁷ It is shown that the introduction of *ex post* flexibility can reverse the direction of the firm's response to an increase in uncertainty.⁸

Denote the cumulative distribution function of θ by $F(\theta, s)$, where s is a parameter representing the dispersion of the distribution. An increase in s represents a mean-preserving spread of the distribution if the following conditions are met:

⁶ This result is not the result of risk aversion. A risk-neutral firm would also select a larger capacity when flexibility is present.

⁷ It is easy to show that an increase in uncertainty reduces expected utility if D is a convex function of θ ($D_{\theta\theta} \geq 0$).

⁸ When *ex post* flexibility is present, an increase in risk aversion increases the optimal capacity if the input is risk reducing ($-D_{x\theta} > 0$). However, if the input is risk increasing ($-D_{x\theta} < 0$), then the effect of increased risk aversion is indeterminate. These results conform exactly to those obtained when *ex post* flexibility is absent (see Hiebert 1989).

⁵ An example is fire retardant material installed in the walls of a building.

$$(12) \quad \int_0^{\infty} F_s(\theta, s) d\theta = 0, \quad \text{and}$$

$$(13) \quad T(y, s) = \int_0^y F_s(\theta, s) d\theta \geq 0 \quad \text{for all } y.$$

The effect of an increase in uncertainty is given by

$$(14) \quad \partial k^* / \partial s = -(\partial^2 EU / \partial k \partial s) / (\partial^2 EU / \partial k^2).$$

Because the second-order conditions imply that the denominator is negative, the sign of (14) is the same as the sign of $\partial^2 EU / \partial k \partial s$.

First, consider the case where no flexibility is present. Differentiating (9) with respect to s gives

$$(15) \quad \partial^2 EU / \partial k \partial s = p \int_0^{\infty} U'(\pi_1)(-D_x(k, \theta) - w'(k) - r) dF_s(\theta, s).$$

Integrating the right-hand side of (15) twice by parts yields

$$(16) \quad \partial^2 EU / \partial k \partial s = p \int_0^{\infty} \partial^2 (U'(\pi_1)(-D_x(k, \theta) - w'(k) - r)) / \partial \theta^2 T d\theta$$

because $F_s(0, s) = F_s(\infty, s) = T(0, s) = T(\infty, s) = 0$. A condition sufficient to determine the effect of increased uncertainty is that $\partial^2 (U'(\pi_1)(-D_x - w' - r)) / \partial \theta^2$ be uniformly positive or negative. If $U'(\pi_1)(-D_x - w' - r)$ is a strictly convex (concave) function of θ , then the firm responds to an increase in uncertainty by increasing (decreasing) the optimal level of capacity.⁹ These conditions are the familiar Rothschild-Stiglitz conditions for signing the effect of a mean-preserving increase in uncertainty.

However, if *ex post* flexibility is present, the Rothschild-Stiglitz conditions are no longer sufficient to determine the firm's response to an increase in uncertainty. Suppose x is more effective in more severe states. Differentiating the first-order condition (5) with respect to s and integrating by parts twice gives

$$(17) \quad \begin{aligned} \partial^2 EU / \partial x \partial s = & p U'(\pi_1(\theta^*)) (-D_{x\theta}(k, \theta^*)) \\ & \cdot T(\theta^*, s) + p \int_0^{\infty} \partial^2 (U'(\pi_1) \cdot (-r)) / \partial \theta^2 T d\theta \\ & + p \int_{\theta^*}^{\infty} \partial^2 (U'(\pi_1)(-D_x(k, \theta) - w'(k))) / \partial \theta^2 T d\theta, \end{aligned}$$

where $\pi_1(\theta^*)$ denotes profit evaluated at θ^* . The second and third terms in (17) are analogous to (16). The first term in (17) arises only when *ex post* flexibility is present and can be called a flexibility effect.¹⁰ Since $-D_{x\theta} > 0$, this effect is always positive and thus tends to increase the firm's optimal capacity commitment in the face of increased uncertainty.

If $U'(\pi_1)(-D_x - w' - r)$ is a concave function of θ , then (16) and (17) can clearly have opposing signs. Hence, it is possible that an increase in uncertainty will cause a firm to enlarge its capacity commitment if *ex post* flexibility is present but not if such flexibility is absent.

The result is similar if x is more productive in less severe states. Differentiating (7) with respect to s and integrating two times by parts yields

$$(18) \quad \begin{aligned} \partial^2 EU / \partial x \partial s = & -p U'(\pi_1(\theta^*)) (-D_{x\theta}(k, \theta^*)) \\ & \cdot T(\theta^*, s) + p \int_0^{\infty} \partial^2 (U'(\pi_1) \cdot (-r)) / \partial \theta^2 T d\theta \\ & + p \int_0^{\theta^*} \partial^2 (U'(\pi_1)(-D_x(k, \theta) - w'(k))) / \partial \theta^2 T d\theta. \end{aligned}$$

The first term is the flexibility effect, which again is positive since $-D_{x\theta} < 0$. Hence, the presence of a positive flexibility term does not depend on the risk-affecting character of the technology. As in the previous case, the introduction of utilization flexibility can reverse the direction of the firm's *ex ante* response to an increase in uncertainty.

An Example

The following example illustrates the results of this paper.

⁹ If the firm is risk neutral, the comparative static results depend on whether $-D_x$ is convex or concave in θ . As a special case, if the firm is risk neutral and θ enters the loss function multiplicatively, then the capacity decision of the firm is unaffected by an increase in risk.

¹⁰ A flexibility effect does not arise in the models of *ex post* choice used by Turnovsky, Hartman, Epstein, and Robison and Barry. Hence, in these models, the Rothschild-Stiglitz conditions are sufficient to determine the impact of an increase in uncertainty.

(a) There are only three possible loss states of the world: $\theta \in \{1/4, 1, 7/4\}$ with probabilities $\alpha/2, 1 - \alpha, \alpha/2$. An increase in α corresponds to a mean-preserving spread of this distribution of θ .

(b) The damage function is given by $(D_0 - 2\sqrt{x})\theta$. For this specification, $-D_x = \theta/\sqrt{x}$ and $-D_{x\theta} = 1/\sqrt{x} > 0$.

(c) The marginal cost of utilizing capacity is a constant, w .

(d) The firm is risk neutral.

When there is no *ex post* flexibility, the first-order condition for optimal capacity choice is

$$\partial E\pi/\partial k = -r - w + p/\sqrt{k} = 0.$$

Let $r = 1/10$, $w = 1/2$, and $p = 6/10$. Then the optimal capacity choice is 1. [Observe that the *ex post* marginal net return to the input $(-D_x - w)$ is negative if the realized value of θ is $1/4$].

On the other hand, when *ex post* flexibility is present, optimal capacity choice is determined by

$$\begin{aligned} \partial E\pi/\partial k &= -r + p(1 - \alpha)(1/\sqrt{k} - w) \\ &+ p(\alpha/2)(7/4\sqrt{k} - w) = 0. \end{aligned}$$

For $\alpha = 1/3$, it can be verified that the optimal capacity choice is $23/14$.

Now let $\alpha = 1/2$. When *ex post* flexibility is absent, the optimal capacity choice is unaffected by a mean-preserving increase in uncertainty. However, when *ex post* flexibility is present, the optimal capacity is $45/26 > 23/14$. An increase in uncertainty increases the demand for capacity.¹¹ The introduction of *ex post* flexibility qualitatively alters the firm's response to an increase in uncertainty.

Conclusions

The existing literature has assumed that all damage control decisions must be made *ex ante*. This

paper has examined the implications of *ex post* utilization flexibility for the firm's choice of capacity to supply a loss-reducing input. It was shown that the optimal capacity is always greater when utilization is flexible rather than fixed. In addition, the total effect of increased uncertainty on the optimal choice includes a flexibility effect when *ex post* flexibility is present. This flexibility effect always tends to increase the optimal capacity.

[Received November 1989; final revision received June 1990.]

References

- Centner, T. J., and M. E. Wetzstein. "Reducing Moral Hazard Associated with Implied Warranties of Animal Health." *Amer. J. Agr. Econ.* 69(1987):143-50.
- Epstein, L. "Production Flexibility and the Behavior of the Competitive Firm Under Price Uncertainty." *Rev. Econ. Stud.* 45(1978):251-61.
- Feder, G. "Pesticides, Information, and Pest Management Under Uncertainty." *Amer. J. Agr. Econ.* 61(1979):97-103.
- Hartman, R. "Factor Demand with Output Price Uncertainty." *Amer. Econ. Rev.* 66(1976):675-81.
- Hiebert, L. D. "Optimal Loss Reduction and Increases in Risk Aversion." *J. Risk and Insurance* 56(1989):300-305.
- . Hiebert, L. D. "Self Insurance, Self Protection and the Theory of the Competitive Firm." *S. Econ. J.* 50(1983):160-68.
- Just, R. E., and R. D. Pope. "Production Function Estimation and Related Risk Considerations." *Amer. J. Agr. Econ.* 61(1979):276-84.
- Lichtenberg, E., and D. Zilberman. "The Econometrics of Damage Control: Why Specification Matters." *Amer. J. Agr. Econ.* 68(1986):261-73.
- Robison, L. J., and P. J. Barry. *The Competitive Firm's Response to Risk*. New York: MacMillan Publishing Co., 1987.
- Rothschild, M., and J. E. Stiglitz. "Increasing Risk: I, A Definition." *J. Econ. Theory* 2(1970):225-43.
- Turnovsky, S. J. "Production Flexibility, Price Uncertainty and the Behavior of the Competitive Firm." *Int. Econ. Rev.* 14(1973):395-413.

¹¹ In this example only the flexibility effect is present.

The Application and Economic Interpretation of Selectivity Models

Chung L. Huang, Robert Raunikaar, and Sukant Misra

Substantial differences in economic interpretations may be inferred from the estimated selectivity model results. Using the demand for frozen concentrated orange juice (FCOJ) as an example, the study suggests that the consumption patterns are quite similar between national brand and private label when the marginal effects and elasticities are evaluated based on actual observations rather than unobserved latent variables. Results show that private label FCOJ is considered as a normal (inferior) good if the income effect is assessed based on actual (potential) purchases. The estimated cents-off elasticities for private label FCOJ also vary substantially depending on which approach is used.

Key words: conditional expectation, elasticity, marginal effect, orange juice, selectivity bias, switching regression, unconditional expectation.

The modeling and empirical measurement of consumer preferences and behavior have been of great importance to demand analysts and marketing researchers. A common approach to these modeling efforts is to assume that various measures of individual behavior are determined in a utility maximization framework in which an individual selects one element from a discrete set of decision alternatives. Examples of such choice problems include decisions on travel mode, labor force participation, educational level, and brands of commodity purchased.

A frequently encountered problem in marketing research is the quantification of individual demand which may be influenced by unobserved behavior or preference; for example, a consumer may or may not purchase a particular product during the period of the survey. Also, individual-based data generally require researchers to deal with potentially biased estimates arising from the selection process that generated the sample. Recognizing this problem of individual self-selection, Heckman developed a procedure which transforms the problem from one of missing data on the dependent variable (no earning data for nonparticipants in the labor force) to one of specification error or

omitted variable resulting from sample selection bias. Following Heckman's study of labor supply, many econometric models have been developed to generate and estimate the conditional expectation of actions not taken. These models are generally called selectivity models.

Dolton and Makepeace examine the structure of a two-sector selectivity model with a probit selector equation. Assuming that selectivity bias is statistically significant, they investigate the appropriate interpretation of the sign of the sample selection coefficient. Dolton and Makepeace conclude that the notion of a correct sign for the selectivity coefficient is difficult to resolve on economic grounds and that any inferences from the results must be clearly qualified.

Many studies have extended the basic sample selection framework into a general class of econometric applications known as switching regression models. Empirical examples include demand for housing (Lee and Trost), brand name selection and orange juice consumption (Lee, Brown, and Schwartz), coupon promotional programs and demand for orange juice (Lee and Brown 1985), and at-home and away-from-home food expenditure patterns (Lee and Brown 1986). Switching regression models have provoked some interesting debate in econometric literature. Poirier and Raud argue that inherent confusion in the endogenous switching model specifications can cause ambiguous interpretation problems. However, as Maddala points out, the observational equivalence in model specifications

Chung L. Huang and Robert Raunikaar are professors and Sukant Misra is a postdoctoral associate, Department of Agricultural Economics, Georgia Experiment Station, University of Georgia.

The authors wish to thank Dale H. Carley, Joseph C. Purcell, and the *Journal* referees for helpful comments on earlier drafts of this manuscript.

which Poirier and Ruud observed essentially is two different ways of writing the same model (p. 285). He concludes that there is no ambiguity of inferences arising from endogenous switching models and that they are important and useful tools for many economic research applications. Nevertheless, the appropriate interpretation of coefficient estimates requires further exploration.

The primary objective of this study is to focus on proper analysis and interpretation of the empirical results from selectivity models. Similar to McDonald and Moffitt's analysis of the tobit model, this study explores the importance of economic interpretations of coefficient estimates. Specifically, the study will demonstrate how different approaches used for assessing marginal effects and elasticities from selectivity models may produce substantial differences in the interpretation of results. The following sections develop a generalized model for deriving conditional and unconditional marginal effects of the explanatory variables specified. An example from the literature is used to illustrate how qualitative interpretations may change when the marginal effects of the estimated model are evaluated with the proposed approach.

Selectivity Model Specification

The basic structure of a general selectivity model involving two regimes is considered. Given a sample of size T , let I_t^* be a $T \times 1$ column vector of an unobserved latent variable that separates the sample into two groups or regimes. Furthermore, it is assumed that $q_t^{(1)*}$ and $q_t^{(2)*}$ are $T \times 1$ column vectors that represent the quantity demanded for products 1 and 2 in regimes 1 and 2, respectively. Thus, the general model that consists of two regimes which describe consumer's demand for products 1 and 2 can be specified by a set of simultaneous equations such as

$$\begin{aligned} (1) \quad & I_t^* = S_{tk}\tau_k + \epsilon_t, \quad \text{for } t = 1, 2, \dots, T; \text{ and } k = 1, 2, \dots, K; \\ (2) \quad & q_t^{(1)*} = X_{tm}^{(1)}\alpha_m + u_t, \quad \text{iff } \epsilon_t > -S_{tk}\tau_k, \quad \text{for } m = 1, 2, \dots, M; \text{ and} \\ (3) \quad & q_t^{(2)*} = X_{tn}^{(2)}\beta_n + v_t, \quad \text{iff } \epsilon_t \geq S_{tk}\tau_k, \quad \text{for } n = 1, 2, \dots, N. \end{aligned}$$

S_{tk} , $X_{tm}^{(1)}$ and $X_{tn}^{(2)}$ are matrices of exogenous variables with T observations and the number of variables equals K , M , and N , respectively; τ_k , α_m , and β_n are parameter vectors with dimensions of $K \times 1$, $M \times 1$ and $N \times 1$, respectively. The residuals ϵ_t are normally distributed with

mean 0 and variance 1, and u_t and v_t are normally distributed with mean 0 and variances σ_u^2 and σ_v^2 , respectively. Note that $q_t^{(1)*}$ and $q_t^{(2)*}$ are observed conditional on I_t^* . Because the indicator I_t^* in the selection equation (1) is not observed, the partial samples of $q^{(1)}$ and $q^{(2)}$ are observed in terms of a binary variable, I , such that

$$\begin{aligned} (1.1) \quad & I = 1, \quad \text{iff } \epsilon > -S\tau_k, \\ & = 0, \quad \text{iff } \epsilon \geq S\tau_k. \end{aligned}$$

All subscripts, denoting observation t and/or variable k in the sample, have been omitted to simplify the presentation. The specification of equation (1.1) implies the familiar probit model

$$\text{Prob}(I = 1) = \Phi(S\tau_k), \text{ and}$$

$$\text{Prob}(I = 0) = 1 - \Phi(S\tau_k),$$

where $\Phi(\cdot)$ denotes the standard normal distribution function. Based on equation (1.1), the observed conditional demand equation in each regime can be rewritten as

$$(2.1) \quad q^{(1)} = X^{(1)}\alpha_m + u, \quad \text{iff } I = 1, \text{ and}$$

$$(3.1) \quad q^{(2)} = X^{(2)}\beta_n + v, \quad \text{iff } I = 0.$$

Given that in practice $q^{(1)}$ is observed if and only if $I = 1$, and $q^{(2)}$ is observed if and only if $I = 0$, there is a sample selectivity problem in the observed data. The problem arises because the model implies that the unobserved disturbance ϵ of the selection criterion equation is correlated with the unobserved disturbances of u and v of the demand equations (2.1) and (3.1). Following Maddala, the conditional expected demands for $q^{(1)}$ and $q^{(2)}$ are derived as

$$(2.2) \quad E(q^{(1)}|I = 1) = X^{(1)}\alpha_m - w^{(1)}\sigma_{u\epsilon}, \quad \text{and}$$

$$(3.2) \quad E(q^{(2)}|I = 0) = X^{(2)}\beta_n + w^{(2)}\sigma_{v\epsilon},$$

where $\sigma_{u\epsilon}$ and $\sigma_{v\epsilon}$ are the $\text{cov}(u, \epsilon)$ and $\text{cov}(v, \epsilon)$, respectively; $w^{(1)} = \phi(S\tau_k)/\Phi(S\tau_k)$, and $w^{(2)} = \phi(S\tau_k)/[1 - \Phi(S\tau_k)]$; and $\phi(\cdot)$ is the standard normal density function.

For estimation purposes, equations (2.2) and (3.2) are specified as

$$(2.3) \quad q^{(1)} = X^{(1)}\alpha_m - w^{(1)}\sigma_{ve} + \xi^{(1)}, \quad \text{and}$$

$$(3.3) \quad q^{(2)} = X^{(2)}\beta_n + w^{(2)}\sigma_{ve} + \xi^{(2)},$$

where $\xi^{(1)}$ and $\xi^{(2)}$ are new residuals with zero conditional means and heteroscedastic variances. In essence, the above specified model requires two steps in the estimation process. First, the probability of sample entry is estimated from equation (1.1) by a probit procedure; second, after estimates of τ_k , $w^{(1)}$ and $w^{(2)}$ are computed, they are added to equations (2.3) and (3.3) as omitted variables to account for sample selection bias. Given that the variance of the residuals is heteroscedastic in nature, equations (2.3) and (3.3) should be estimated by weighted least squares. In what follows, it can be shown that the conventional interpretation of marginal effects based on estimated coefficients is not strictly appropriate if demand equations include a subset of independent variables that also appear in the specification of the selection equation.

Conditional and Unconditional Marginal Effects

As Maddala points out, the marginal effects from the model can be assessed in three different ways, depending on the purpose of the research and the desired information (pp. 159–60). The first measurement of marginal effects can be derived in terms of the latent variables such as $q_i^{(1)*}$ and $q_i^{(2)*}$, which represent desired or potential quantities. If the objective is to predict the mean of potential quantities demanded, then the estimated coefficients, α and β , obtained from equations (2) and (3) would represent the appropriate marginal effects. In addition, two other measurements of marginal effects, conditional and unconditional, can be derived from the estimated model based on observed or actual quantities demanded. The following analysis focuses on obtaining predictions from the estimated model based on observed rather than unobserved potential quantities.

For simplicity, let z be the exogenous variable such that $z \in (S \cap X^{(1)} \cap X^{(2)})$. In other words, z represents a common explanatory variable which is included in the specification of equations (1.1), (2.3), and (3.3). By taking partial derivatives of equations (2.2) and (3.2) with respect to z , one obtains the conditional marginal effects of z on $q^{(1)}$ and $q^{(2)}$ as

$$(4) \quad \partial E(q^{(1)}|I=1)/\partial z = \alpha - \sigma_{ue}(\partial w^{(1)}/\partial z),$$

$$= \alpha + \tau\sigma_{ue}[(S\tau_k)w^{(1)} + (w^{(1)})^2], \quad \text{and}$$

$$(5) \quad \partial E(q^{(2)}|I=0)/\partial z$$

$$= \beta - \tau\sigma_{ue}[(S\tau_k)w^{(2)} - (w^{(2)})^2],$$

where α , β , and τ are parameters associated with the z variable in the demand and selection equations (Poirier and Ruud). As shown in equations (4) and (5), α and β alone account for only a portion of the marginal effects derived from equations (2.3) and (3.3). Thus, any interpretation of marginal effects based primarily on the α s and β s may be misleading because the effects of the second terms in equations (4) and (5) were not evaluated (Dolton and Makepeace). As noted earlier, equations (4) and (5) represent the marginal effects that are conditional on partial sample observations specific to each regime. Additional information can be derived based on all observations with respect to each product. Specifically, the unconditional expected values, $E(q^{(1)})$ and $E(q^{(2)})$, can be written as the probability of a regime being selected multiplied by the conditional expected value associated with the selected regime. Thus,

$$(6) \quad E(q^{(1)}) = \Phi(S\tau_k)E(q^{(1)}|I=1), \quad \text{and}$$

$$(7) \quad E(q^{(2)}) = [1 - \Phi(S\tau_k)]E(q^{(2)}|I=0).$$

Following McDonald and Moffitt, it can be seen that the marginal effects based on the total sample for a product can be decomposed into two parts:

$$(6.1) \quad \begin{aligned} \partial E(q^{(1)})/\partial z &= \Phi(S\tau_k)[\partial E(q^{(1)}|I=1)/\partial z] \\ &\quad + E(q^{(1)}|I=1)[\partial \Phi(S\tau_k)/\partial z], \\ &= \alpha\Phi(.) + \tau\phi(.)[X^{(1)}\alpha_m + (S\tau_k)\sigma_{ue}]. \end{aligned}$$

The first part represents the change in the observed $q^{(1)}$ weighted by the probability that regime 1 is selected; the second part represents the marginal probability, $\tau\phi(.)$, or the change in the probability of regime 1 being selected weighted by the expected value of $q^{(1)}$ if regime 1 is selected. Similarly, the marginal effects based on the total sample for product 2 in regime 2 can be shown as

$$(7.1) \quad \begin{aligned} \partial E(q^{(2)})/\partial z &= \beta[1 - \Phi(.)] \\ &\quad - \tau\phi(.)[X^{(2)}\beta_n + (S\tau_k)\sigma_{ve}]. \end{aligned}$$

In addition, the total unconditional marginal effect can be obtained from equations (6.1) and (7.1). Let q be the total demand for both $q^{(1)}$ and

$q^{(2)}$ regardless of whether it is observed for regime 1 or 2. By combining both equations (6.1) and (7.1), the effect of a change in the z variable on total demand for both products can be written as

$$(8) \quad \partial E(q)/\partial z = \alpha\Phi(.) + \beta[1 - \Phi(.)] \\ + \tau\phi(.)[(X^{(1)}\alpha_m - X^{(2)}\beta_n) \\ + (S\tau_k)(\sigma_{ue} - \sigma_{ve})].$$

The derivations of equations (6.1), (7.1), and (8) provide important economic interpretations for decomposing the unconditional marginal effects. In general, the interpretations of these equations are analogous to those offered by Thraen, Hammond, and Buxton for tobit analysis. Equation (8), for example, provides the basis for evaluating the unconditional marginal effect of an independent variable on the total demand for two products when the variable is included in the specifications of selection equation and the demand equations. The first two terms represent the quantity response due to actual purchase of the product and the last term is an adjustment factor for selectivity bias resulting from entry or exit from the market.

Depending on model specifications, slightly different formulas can be derived to evaluate unconditional marginal effects from estimated results. Necessary adjustments for assessing marginal effects under various model specifications are summarized in table 1. Equations (8)–(13) in table 1 are a guide for appropriate evaluation of the total unconditional marginal effect for any given independent variable on demand for $q^{(1)}$ and $q^{(2)}$. As shown in table 1, the need to adjust the estimated coefficients in the computation of unconditional marginal effects arises

only when an independent variable in the demand equations is included in the sample selection equation. The magnitude of the adjustment factors is the same regardless of whether the variable is included in both or only one of the demand equations. Therefore, the underlying factor determining whether an adjustment is necessary depends on the specification of the demand equation relative to the selection equation.

Empirical applications employing Heckman's procedure or tobit analysis are special cases of the general selectivity model in table 1. Specifically, equation (10) is applicable to a single regime model in which only a selection criterion and a demand equation are specified and Heckman's sample selection procedure is used in the estimation process (e.g., Cheng and Capps). In this case, the conditional marginal effects should be adjusted using equation (5). The unconditional effects can be evaluated and decomposed as shown in equation (10), except for the terms $X^{(1)}\alpha_m$ and σ_{ue} . Similarly, the tobit analysis is a special case of equation (12) in which the sample selection process is embedded in the demand equation as a single equation model. Thus, equation (12) corresponds to the unconditional marginal effect of $\alpha\Phi(.)$ for an explanatory variable, as McDonald and Moffitt demonstrated for tobit analysis.

Application and Interpretation of Results

The data used for the illustration in this study are drawn from a recent study by Lee, Brown, and Schwartz (LBS). They developed a switching regression model similar to that of equations

Table 1. Decomposition of Total Unconditional Marginal Effects for Various Selectivity Model Specifications

Equation	Variable z included			Marginal Effect of a Change in Variable z on Demand	
	(1.1)	(2.3)	(3.3)	Resulting from Actual Purchase	Selectivity Bias Adjustment from Entry or Exit from the Market
				$\alpha\Phi(.) +$	
(8)	*	*	*	$\beta[1 - \Phi(.)]$	$\tau\phi(.)[(X^{(1)}\alpha_m - X^{(2)}\beta_n) + (S\tau_k)(\sigma_{ue} - \sigma_{ve})]$
(9)	*	*		$\alpha\Phi(.)$	$\tau\phi(.)[(X^{(1)}\alpha_m - X^{(2)}\beta_n) + (S\tau_k)(\sigma_{ue} - \sigma_{ve})]$
(10)	*		*	$\beta[1 - \Phi(.)]$	$\tau\phi(.)[(X^{(1)}\alpha_m - X^{(2)}\beta_n) + (S\tau_k)(\sigma_{ue} - \sigma_{ve})]$
(11)		*	*	$\alpha\Phi(.) +$	
				$\beta[1 - \Phi(.)]$	
(12)		*		$\alpha\Phi(.)$	
(13)			*	$\beta[1 - \Phi(.)]$	

(1.1), (2.3), and (3.3) to examine the demand for national brand and private label frozen concentrated orange juice (FCOJ) in the United States. They used a two-stage estimation procedure for studying the joint determination of the self-selection process of brand labels and the demand for FCOJ. LBS reported that the demand for national brand frozen concentrated orange juice is more price responsive than that for the private label. They also found that socioeconomic and demographic characteristics generally have opposite impacts on the demand for these two products. Furthermore, LBS concluded that while national brand FCOJ is a normal good, the private label FCOJ is an inferior good and the consumption patterns are quite different between the two brands of FCOJ.

As previously discussed, marginal effects obtained from switching regression models may be assessed in terms of unobserved potential or observed actual quantities. LBS interpret their results on the basis of desired or potential effects. Thus, their interpretations of the results based on the estimated regression coefficients may be appropriate because the study's purposes and interests were focused on the latent variables of the model. However, for our study, the alternative interpretations based on observed quantities demanded are discussed using the LBS study. The analysis extends the LBS results to illustrate that additional information and different interpretations can be derived from the estimated results if the marginal effects are evaluated in terms of actual observations instead of unobserved latent variables.

Results of the LBS study are partially reproduced in table 2. The conditional and unconditional marginal effects were computed using equations (4), (5), and (8) based on the sample information and estimated parameters reported by LBS. Depending on the magnitude of the adjustment, the sign on marginal effect may differ from the estimated coefficient. For example, after selection bias was considered, the results suggest that female employment, female education, and female age all have negative rather than positive effects on per capita demand for private brand FCOJ (table 2). In general, the signs on marginal effects for national brand FCOJ were not changed, but the coefficients tend to be smaller relative to the LBS unadjusted estimates. For private label FCOJ, the signs on marginal effects in most cases were changed and these signs along with the size of the marginal effects become more consistent with those of national brand FCOJ.

From the perspective of commodity demand analysis, it is of interest to determine how various socioeconomic and demographic factors affect demand for FCOJ regardless of brand labels. To this end, the corresponding unconditional marginal effects are presented in table 2. Except for inherent conceptual differences, the interpretation of unconditional marginal effects is quite straightforward and does not require further elaboration.

As shown in table 2, this study suggests that a similar relationship exists between per capita demand for national brand FCOJ and private label FCOJ except for the presence of children and

Table 2. Estimated Conditional and Unconditional Marginal Effects of Selected Explanatory Variables in the Demand for Frozen Concentrated Orange Juice

Explanatory Variables	Regression Coefficients			Conditional Effects		Unconditional ^a
	Probit Model	National Brand	Private Label	National Brand ^a	Private Label ^a	
Cents-off	.0061	.3975	.1903	.3647	.4254	.3811
Female employment	-.0031	-.0875	.0489	-.0708	-.0706	-.0690
Female education	-.0630	-1.6176	1.6819	-1.2792	-.7457	-1.0678
Female age	-.0049	-.1439	.0766	-.1176	-.1122	-.1123
Income	.0110	.3918	-.0949	.3327	.3290	.3251
Children	-.0261	.9888	-.9292	1.1290	-1.9350	.1385
Family size	.0082	-1.4702	-.9783	-1.5142	-.6623	-1.2394
Winter	.0236	2.2763	.0711	2.1495	.9805	1.7522
Spring	.0405	.8294	-.0633	.6119	1.4973	.8791
Summer	-.5658	-.5658	.7069	-.4149	-.3759	-.3859
Atlantic region	.0826	2.3971	-1.0045	1.9534	2.1784	1.9795
Central region	-.0774	-5.2288	2.5073	-4.8131	-.4752	-3.3447
Southern region	-.0740	-3.8270	3.4680	-3.4235	.6165	-2.0589

Source: The estimated demand parameters reproduced here are obtained from Lee, Brown, and Schwartz (table 4, p. 6). Only those variables that require adjustments in the computation of relevant marginal effects are presented.

^a Computed from equations (4), (5), and (8), respectively.

for households that reside in the southern region of the United States. Furthermore, if the income effect is based on actual rather than potential purchases, private label FCOJ is considered as a normal good rather than an inferior good, as LBS suggest. After adjusting for selectivity bias, the results show that the income effects based on actual purchases of national brand and private label FCOJ are .3327 and .3290, respectively. The consistency of the income effects suggests the demand for FCOJ reflects primarily differences in shopping behaviors with respect to brand name selection.

LBS reported that the elasticities for cents-off for national brand and private label FCOJ are .36 and .09, respectively. If, indeed, potential buyers of private label FCOJ are highly unresponsive to special cents-off price discount, an advertising campaign in the form of price discount is clearly not an effective method to promote the sale of private label FCOJ. However, if the marginal effects are based on actual purchases, the elasticities for cents-off are .33 for national brand FCOJ and .39 for private label FCOJ. Again, the analysis shows that consumers' response to special price discount in actual purchases are very similar for both national brand and private label, suggesting the effectiveness of special cents-off on FCOJ is about the same regardless of brand names.

Conclusions

The purposes of this study are twofold. First, the study calls attention to different approaches that may be used for analyzing the estimated selectivity model results. Second, the study demonstrates how qualitative interpretations may differ depending on whether the marginal analysis was based on potential or actual impacts. Furthermore, when an independent variable is included in the specification of a selection equation and the conditional demand equation, some adjustments may be required in order to appropriately assess the economic implications suggested by the data.

Using the demand for FCOJ as an example, the study suggests that the consumption patterns are quite similar between national brand and private label when the marginal effects are evaluated in terms of actual observations instead of unobserved latent variables. Specifically, the results show that private label FCOJ is considered

as a normal (inferior) good if the income effect is assessed based on actual (potential) purchases. The estimated cents-off price elasticities for private label FCOJ also vary substantially depending which approach is used in the analysis.

Although correction for self-selectivity bias is imperative for obtaining unbiased coefficient estimates, the study demonstrates the importance of distinguishing the different approaches upon which assessment of the empirical results and interpretations are based. The same approach can be applied to other forms of selectivity models, such as the popular Heckman sample selection procedure.

[Received July 1989; final revision received May 1990.]

References

- Cheng, H. T., and O. Capps. "Demand Analysis of Fresh and Frozen Finfish and Shellfish in the United States." *Amer. J. Agr. Econ.* 70(1988):533-42.
- Dolton, P. J., and G. H. Makepeace. "Interpreting Sample Selection Effects." *Econ. Letters* 24(1987):373-79.
- Heckman, J. "The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimation for Such Models." *Ann. Econ. and Soc. Measure.* 5(1976):475-92.
- Lee, J. Y., and M. Brown. "Coupon Redemption and the Demand for Frozen Concentrated Orange Juice: A Switching Regression Analysis." *Amer. J. Agr. Econ.* 67(1985):647-53.
- . "Food Expenditures At Home and Away from Home in the United States—A Switching Regression Analysis." *Rev. Econ. and Statist.* 68(1986):142-47.
- Lee, J. Y., M. Brown, and B. Schwartz. "The Demand for National Brand and Private Label Frozen Concentrated Orange Juice: A Switching Regression Analysis." *West. J. Agr. Econ.* 11(1986):1-7.
- Lee, L. F., and R. P. Trost. "Estimation of Some Limited Dependent Variables Models with Applications to Housing Demand." *J. Econometrics* 8(1978):357-82.
- Maddala, G. S. *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge: Cambridge University Press, 1983.
- McDonald, J. F., and R. A. Moffitt. "The Use of Tobit Analysis." *Rev. Econ. and Statist.* 60(1980):318-21.
- Poirier, D. J., and P. A. Ruud. "On the Appropriateness of Endogenous Switching." *J. Econometrics* 16(1981):249-56.
- Thraen, C. S., J. W. Hammond, and B. M. Buxton. "Estimating Components of Demand Elasticities from Cross-Sectional Data." *Amer. J. Agr. Econ.* 60(1978):674-77.

Erratum

Corrected materials are presented below for errors that occurred in an article entitled "The Demand for Canadian Fats and Oils: A Case Study of Advertising Effectiveness," by Ellen W. Goddard and Alex K. Amuah, in the August 1989 issue of this *Journal*, pages 741-49 (*Amer. J. Agr. Econ.* 71(1989):741-49).

Advertising Elasticities

The advertising elasticity table (p. 747) is erroneous. The corrected advertising elasticities are as follows:

Table 6. Advertising Elasticities

	Butter	Margarine	Shortening	Vegetable Oils
Mean				
Butter	.005	-.06	-.02	.003
Margarine	-.01	.04	.01	-.05
Shortening	-.04	.14	.03	.07
Vegetable oils	.04	.06	.02	.07
1986:4				
Butter	.001	-.05	-.02	-.02
Margarine	.0002	.05	.02	.01
Shortening	-.01	.08	.03	.05
Vegetables oils	.02	-.01	-.01	.0001

The text of the paper beginning with the first new paragraph, column 2, page 747 should be corrected as follows:

Regarding the cross-advertising effects, margarine advertising has had a negative effect on butter demand. On average, a 1% increase in the advertising of margarine leads to a 0.06% reduction in the demand for butter. The effect of butter cross-advertising on margarine is smaller than the cross-effect from margarine on butter on average. A 1% increase in the advertising of butter leads to a 0.01% decrease in the demand for margarine on average.

Vegetable oil advertising has a slight positive effect on butter demand. On average a 1% increase in the advertising expenditures of shortening leads to a 0.02% decrease in the demand for butter.

On average, margarine demand responds positively to shortening advertising. A 1% increase in the advertising of shortening leads to a 0.01% increase in the demand for margarine. Margarine responds negatively to vegetable oil advertising. A 1% increase in the advertising of vegetable oil leads to a 0.05% reduction in the demand for margarine.

Butter advertising has a negative effect on shortening and margarine demand. A 1% increase in the advertising of butter leads to a .04% decrease in the demand for shortening on average. Margarine advertising does not have a negative effect on shortening demand. A 1% increase in margarine advertising leads to a .14% increase in the demand for shortening on average.

A 1% increase in the advertising of butter leads to a .04% increase in vegetable oil demand. A 1% increase in the advertising of margarine leads to a .06% increase in vegetable oil demand. A 1% increase in the advertising of shortening leads to a 0.02% increase in the demand for vegetable oil.

The Impact of Increased LDC Food Production on LDC Food Imports: Comment

Patrick J. Gormely

If a less developed country (LDC) increases its food production, will its food imports decrease or increase? This question is of obvious importance not only to those concerned with development of LDCs but also to agricultural exporters in the United States and elsewhere.¹

In a series of papers (beginning with one in this *Journal*) de Janvry and Sadoulet (hereafter JS) have used simulation models to examine the impact of increased LDC food production on LDC food imports (1986, 1988; De Janvry, Sadoulet, and White). I offer a critique of their model after presenting a brief background to the debate.

Two Answers to the Question

Two opposing views exist. On one side are those who argue that increased food production will lead to decreased food imports. Examples are provided by the LDC "food gap" ("food crisis") literature that projects LDC food production, food consumption, and food imports. All the projections of which I am aware show that the more rapid the growth of food production, the slower the growth of food imports. In these projection models, the stimulus imparted to LDC economic development by a given increase in LDC food production is not large enough to cause an equivalent increase in food consumption, so food imports decrease.² Similar views have been expressed

by food exporters in the United States, who are worried that increases in LDC food production (perhaps caused by U.S. assistance to LDC agriculture) will lead to reduced LDC food imports. These groups, too, have in mind models of development wherein increases in LDC food production do not lead to equivalent increases in food consumption, so food imports decrease.

In striking contrast to these views, some economists have argued that increases in LDC food production will lead to increased LDC food imports.³ Mellor has long argued that exporters of food to LDCs will benefit from increases in LDC food production: "Perhaps surprisingly, policies which stimulate development of the domestic agricultural sectors of these countries are likely to provide the most rapid growth in their agricultural imports" (1978, p. 38). And, more recently, Mellor has argued that "It is one of the paradoxes of Third World Development that, as the rate of technological change in food production increases, many developing countries find demand increasing even more rapidly, with consequent increases in food imports. Bachman and Paulino (1979) show that rapid rates of Third World agricultural growth are associated with increases in food imports" (1988b, pp. 424-25).

Others who have recently argued that increased LDC food production will lead to increased LDC food imports include Houck, Kellogg, Paarlberg, and Purcell. These authors have in mind models of development wherein the stimulus imparted to LDC economic development by a given increase in LDC food production is large enough to cause more than an equivalent increase in LDC food consumption, so food imports increase.⁴

Patrick J. Gormely is an associate professor, Department of Economics, Kansas State University.

¹ Sometimes the issue posed is the impact of U.S. assistance to LDC agriculture on LDC agricultural imports from the U.S. This formulation introduces two additional, separate issues: the impact of U.S. assistance on LDC agricultural development and the U.S. share of LDC agricultural imports. These two separate issues are omitted here in order to focus on the central issue: the impact of increased LDC agricultural production (however caused) on LDC agricultural imports (from whatever source).

² FAO (1981) projects LDC food production and consumption to the year 2000, assuming both an "ambitious improvement" over past trends (scenario A) and a "modest improvement" over trends (scenario B). For 90 developing countries together, in the year 2000, scenario A's cereal production exceeds B's by 90 million metric tons (MT), but cereal consumption is higher by only 49 million MT, so cereal imports are 41 million MT lower.

Paulino also projects the food gap for year 2000. He says that "the potential food gap can ultimately be closed by accelerating [food] output growth in developing countries . . ." (p. 56). Also see IBRD (pp. 73-74).

³ Those so arguing share the hope that there is a harmony between the economic interests of U.S. and LDC agricultural producers.

⁴ I think these arguments are erroneous, but this is not the place for a detailed criticism. One point here will suffice. Mellor cites (as do many others) the Bachman-Paulino finding that the sixteen countries experiencing (during 1963-75) rapid growth of food production (58%) increased their food imports by 133%. Overlooked or ignored by all these authors is that Bachman-Paulino also showed (table 23, p. 86) that the 70 countries experiencing slower growth of food production (30%) increased their food imports by 332%. The Bachman-Paulino figures cannot serve as evidence that faster food production growth leads to faster food import growth. Specifically, Mellor's assertion of an association between rapid rates of agricultural growth and increased food imports is misleading.

The de Janvry-Sadoulet Simulation Model and the Interpretation of Its Results

The JS simulation model represents an important improvement in the tools brought to bear on the question of the impact of increased LDC agricultural production on LDC agricultural imports. As I hope is clear, my criticism is directed not against their important modeling effort but rather against their interpretation of its results.

In the JS model the LDC has three sectors: tradable agriculture (using labor and a fixed amount of land as inputs), nontradables (using labor and capital as inputs), and tradable industry (using labor and imperfectly substitutable domestic and imported capital goods as inputs).⁵ The assumption that tradable industrial production uses imported capital goods plays a key role. This assumption makes the quantity of foreign exchange available for importing capital goods a crucial constraint on economic growth. And this in turn means that agricultural growth's essential contribution to the LDC's total economic growth is to allow a reduction of agricultural imports, thereby freeing foreign exchange to pay for additional imported capital goods. How does a model making a reduction of agricultural imports a driving force in development end up predicting an increase in those same agricultural imports? It turns out that one or the other of two questionable assumptions is necessary.

The base-line simulation assumes that the LDC is exporting tradable industrial goods and importing both agricultural goods and capital goods. The model then simulates the impact of a limited technological change in agriculture on the volume of LDC agricultural imports. A 15% increase in land productivity is assumed. The technological change causes agricultural production to increase. The increase in agricultural production exceeds the consequent increase in agricultural consumption, so agricultural imports decrease. (Note that such an import decrease is predicted by both partial and general equilibrium static analysis.) The dynamic JS model now goes to work. The decrease in agricultural imports frees foreign exchange, allowing an increase in the importation of capital goods, thereby causing a subsequent increase in industrial production. After some years the cumulative growth of the economy and the consequent income increases are large enough to cause the increase in agricultural consumption to equal the original limited increase in agricultural production, thereby returning agricultural imports to their original level.⁶ This appears to support the JS conclusion that "tech-

nological change in LDC agriculture can create strong economy-wide growth and income effects with the potential of increasing the level of agricultural imports in future years" (1988 p. 1).

But this conclusion that the LDC's agricultural imports eventually return to and then go on to exceed their original level depends upon the unrealistic assumption of a limited increase in agricultural productivity. The assumption is unrealistic because, if development is constrained by the quantity of imported capital goods, it will presumably be in the LDC's interest to produce substitutes for all imports except the imported capital goods which it cannot produce. This means that the LDC will have an incentive to reduce agricultural imports as much as possible, so one cannot view the agricultural productivity increase as a one-time increase that will eventually be overcome by demand increases. A more realistic simulation is one incorporating a sustained increase in agricultural productivity.

In fact, JS report the result of a sustained increase of 1.4% per year in land productivity, but they give it inadequate attention. JS report that "with a sustained 1.4% productivity growth in agriculture and 0[%] productivity growth in industry, *import demand for food and feed grains falls continuously*, and the country eventually reaches self-sufficiency (1988, p. 21 and fig. 4). [My emphasis added.]

I suggest this simulation result destroys any claim of compatibility between increases in LDC agricultural production and increases in LDC agricultural imports. On the contrary, the JS model shows that sustained productivity growth in agriculture by itself will cause agricultural imports to decrease until they eventually reach zero.⁷

JS (1983) do show that if, in addition to the sustained productivity growth in agriculture, there is a large enough autonomous productivity growth in the industrial tradables sector, the demand for agricultural products will increase enough eventually to cause agricultural imports to return to and then exceed their original level. In effect, JS have shown that sustained agricultural productivity growth is compatible with agricultural import growth only if it is accompanied by sufficiently rapid autonomous growth of the industrial tradables sector. This is analogous to showing that the application of kerosene is compatible with extinguishing a fire if accompanied by the application of sufficiently large amounts of water! The word "compatible" is inappropriate in both arguments.

An Alternative Interpretation

At first glance it appears that the JS model supports the argument that increased LDC food production leads to increased LDC food imports. But the interpreta-

⁵ JS (1986) used a two-sector model. I focus on the latest formulation (1988).

⁶ JS (1988) reports two types of limited (15%) increase in land productivity. In the first the increase takes place all in one year; it takes more than 12 years for agricultural imports to return to their original level. In the second the increase in productivity takes place over 10 years (averaging 1.4% per year); it takes more than 17 years for agricultural imports to return to their original level.

⁷ Of course, productivity increases in agriculture can take the LDC beyond agricultural self-sufficiency, turning the LDC into an exporter of agricultural products.

tion of the JS results presented here shows this impression is wrong. The JS simulation shows an essential incompatibility between sustained LDC agricultural growth and LDC agricultural import growth, an incompatibility which can be prevented from manifesting itself only if there is sufficiently rapid autonomous growth in the nonagricultural parts of the economy. Contrary to first impression, the JS model supports the argument that increased LDC food production leads to decreased LDC food imports.

There has been in recent years a gradual change of emphasis in the argument of writers who see harmony between LDC agricultural growth and increases in LDC agricultural imports. In the past, some proponents of harmony argued or appeared to argue that LDC agricultural growth by itself could stimulate enough LDC economic growth to cause LDC agricultural imports to increase (Mellor 1978, 1983). Recent statements are sometimes more guarded, mentioning "growth arising autonomously in other sectors" (Mellor 1986), "some additional growth initiated outside of agriculture" (Mellor 1988a), and "growth independently generated in the rest of the economy" (Islam) as contributors to the economic growth that eventually causes LDC agricultural imports to increase. The JS model shows that these qualifications to the previous arguments are needed.

[Received March 1988; final revision received July 1990.]

References

- Bachman, Kenneth L., and Leonardo A. Paulino. *Rapid Food Production Growth in Selected Developing Countries*. Washington DC: International Food Policy Research Institute, 1979.
- de Janvry, A., and E. Sadoulet. "The Conditions for Compatibility between Aid and Trade in Agriculture." *Econ. Develop. and Cultur. Change* 37(1988):1-30.
- . "The Conditions for Harmony Between Third World Agricultural Development and U.S. Farm Exports." *Amer. J. Agr. Econ.* 68(1986):1340-46.
- de Janvry, A., E. Sadoulet, and T. K. White. *Foreign Aid's Effect on U.S. Farm Exports*. Washington DC: U.S. Department of Agriculture For. Agr. Econ. Rep. No. 238, 1989.
- Food and Agriculture Organization. *Agriculture: Toward 2000*. Rome, 1981.
- Houck, J. P. "Foreign Agricultural Assistance: Ally or Adversary?" Dep. Agr. and Appl. Econ. Staff Pap. No. P86-50, University of Minnesota, Nov. 1986.
- International Bank for Reconstruction and Development. *Sub-Saharan Africa: From Crisis to Sustainable Growth*. Washington DC, 1989.
- Islam, N. "Third World Food Markets: Option for Agricultural Exporters? An Overview." *IFPRI Policy Briefs* 2. Washington DC: International Food Policy Research Institute, 1988.
- Kellogg, Earl D. *Agricultural Development in Developing Countries and Changes in U.S. Agricultural Exports*. Washington DC: Consortium for International Cooperation in Higher Education, March 1987.
- Mellor, John W. *Agricultural Development in the Third World: The Food, Development, Foreign Assistance, Trade Nexus*. Reprint No. 124. Washington DC: International Food Policy Research Institute, 1988a.
- . "Food Demand in Developing Countries and the Transition of World Agriculture." *Eur. Rev. Agr. Econ.* 15(1988b):419-36.
- . *Food Policy Statement Number 6*. Washington DC: International Food Policy Research Institute, 1986.
- . "Food Prospects for the Developing Countries." *Amer. Econ. Rev.* 73(1983):239-43.
- . "Third World Development and the Demand for Agricultural Exports—The Role of the United States." *World Agricultural Trade: The Potential for Growth*. Kansas City MO: Federal Research Bank of Kansas City, 1978.
- Paarlberg, R. L. "U.S. Agriculture and the Developing World: Partners or Competitors?" *U.S. Agriculture & Third World Development: The Critical Linkage*, ed. R. B. Purcell and E. Morrison. Boulder CO: Lynne Rienner, 1987.
- Paulino, Leonardo A. *Food in the Third World: Past Trends and Projections to 2000*. Washington DC: International Food Policy Research Institute, 1986.
- Purcell, R. B. "Introduction." *U.S. Agriculture & Third World Development: The Critical Linkage*, ed. R. B. Purcell and E. Morrison. Boulder CO: Lynne Rienner, 1987.

The Impact of Increased Food Production on Less Developed Country Food Imports: Reply

Alain de Janvry and Elisabeth Sadoulet

We are pleased that Patrick Gormely correctly restates, in his own language, the same conclusions which we have reached in our analyses of the impact of less developed country (LDC) productivity growth in cereals on their import demand for these same commodities: that, unless productivity growth in cereals is accompanied by productivity growth in other sectors of the economy or by additional income increasing measures, the LDCs cereal imports will decline. As Gormely also correctly points out, the significance of these results is to dismiss statements that have been made by many that conflicts between "aid" (the promotion of productivity growth in LDC cereals production) and "trade" (LDC cereal import demand) simply do not exist. We showed, through general equilibrium modeling, that there is a time path whereby import demand first tends to fall. Whether income effects, including those induced by productivity growth in cereals, eventually accelerate effective demand beyond domestic production depends on a number of structural features and policy interventions which allow the eventual achievement of harmony between aid and trade.

What remains unstated in Gormely's comment are two other fundamental aspects of these results. The first is that most LDCs cannot escape the need to increase productivity growth in their food/feed grains sector as a source of foreign exchange savings if their foreign exchange earning opportunities are not sufficiently buoyant in agroexports, mining, or manufacturing, an unfortunate reality for many late-coming LDCs in today's international economy. For this

reason, green revolution-type aid to LDCs is, perhaps unfortunately from the standpoint of more developed country (MDC) exporters, a necessary component of any meaningful aid program.

The second is that there exists a wide panoply of complementary developmental interventions that can be used to reduce this conflict, at least when calculating the present value of net trade effects with reasonable discount rates. These interventions include assistance to productivity growth in other sectors of the economy, particularly those which are more labor intensive and thus create larger multiplier effects on food demand. Reconciliation through productivity gains in manufacturing is consequently easiest in the newly industrialized countries where those other sectors are already large. Interventions also include the greater opening of MDC markets to LDC manufactures and agroexports as well as initiatives to reduce the debt burden on LDCs. Structural reforms that promote equity and policy reforms that reduce anti-labor biases also increase aggregate effective demand for cereals. Reforms that create income effects have the greatest long-run payoffs in terms of import demand in the lowest income countries since it is where budget shares and income elasticities for staple foods are the highest.

These results suggest that the reconciliation of eventual aid-trade conflicts requires going beyond a project-by-project approach to aid and relocating such interventions within the broader picture of their intersectoral, intertemporal, and interpersonal effects, a vision which has unfortunately been increasingly lost among international development agencies which have too often reduced the difficult task of development to the simpler one of project evaluation.

[Received August 1990; no revision.]

Alain de Janvry is a professor, and Elisabeth Sadoulet is a lecturer, Department of Agricultural and Resource Economics, University of California, Berkeley.

Marketing Order Impacts on Farm–Retail Price Spreads: Comment

Nicholas J. Powers

In a recent article in this *Journal*, Thompson and Lyon (TL) reported that the FOB–retail price spreads for California-Arizona (CA) navel oranges narrowed when marketing order handler prorates were suspended during the 1984/85 season (p. 657). TL's findings are suspect for several reasons.

One, TL's prices are for a heterogeneous commodity, which makes it impossible to accurately differentiate the effect of varying quality from a suspension. The composition of shipment grades and sizes changes over time (and even if they did not change, the proportion of shipments in each grade and size category does). This point is important because when shipments include relatively large volumes of undesirable sizes and grades, the price spread widens because of large shrink.¹

Two, TL's analysis is based on weekly data for the last half of the 1984/85 season and the subsequent two seasons. They did not identify the suspensions during the 1985/86 and 1986/87 seasons. Failure to correctly identify the suspension periods can influence the findings.

Three, TL's price spread specifications are suspect in lieu of the findings by Pick, Karrenbrock, and Carman that weekly price spreads for fresh CA navels display asymmetric response to changes in the FOB price and the lag adjustment to FOB price changes is one to three weeks.² TL also assumed that the marketing margin equations are in reduced form, but retail and FOB prices and shipments are determined simultaneously with handler's supply and the retail demand for navels. Consequently, an instrumental variable estimator is appropriate.

Four, TL's explanation of prorates' influence on price spreads is that producers and handlers favoring prorate maintained that the price spread increased as FOB prices fell while actual retail prices persisted at previous levels. Consumer advocates argued that the

availability of a greater quantity and a wider variety of sizes benefited consumers and narrowed price spreads (p. 649). This explanation is vague.

In this comment, the influence of a suspension on price spreads is analyzed by overcoming the last three problems of TL's article. The findings of this analysis illustrate that TL's findings are sensitive to the time period examined, data, and methods of analysis, and the comment also explains a suspensions' influence on price spreads. However, this analysis, like TL's, is limited by using average prices, which prevents an accurate measurement of the influence of varying product quality.

FOB-Retail Price Spreads

A static marketing margin equation similar to TL's is appropriate because such an equation is consistent with the monthly data of this analysis and the findings by Pick, Karrenbrock, and Carman that lag adjustments occur mostly within the month. The relative model is used to analyze the price spread (Wohlgenant and Mullen). In linear form it is

$$M_t = b_0 + b_1 RP_t + b_2 RP_t Q_t + b_3 MIP_t + b_4 APR_t + b_5 MAY_t + b_6 JUN_t + b_7 SUSP_t + e_t,$$

where the subscript t is the monthly observation, e is the residual term, the b 's denote the estimated intercept and estimated coefficients for the explanatory variables, M is the retail–FOB marketing margin for fresh navel oranges (the difference between the retail and the FOB prices), RP is the retail price, and MIP is the marketing input prices.³ While TL inserted a seasonal binary variable for May, in this analysis binary variables for the last three months of the marketing season (April, May, June) are included to measure seasonal effects. The influence of a suspension on average price spreads is measured by $SUSP$. The binary variable $SUSP$ equals one for a suspension and zero otherwise. In contrast with TL, this analysis identifies all of the suspension periods dur-

The author is an agricultural economist with the Economic Research Service, U.S. Department of Agriculture.

The views expressed herein do not necessarily reflect those of the U.S. Department of Agriculture.

The author acknowledges the helpful comments of an anonymous *Journal* reviewer.

¹ In the produce industry, shrink refers to the difference between what is sold at retail and what is shipped by handlers. Most shrink is the result of product deterioration, with the product ending-up in the garbage bin.

² Two of TL's price spread series are the same as used by Pick, Karrenbrock, and Carman.

³ The marginal cost and markup models were eliminated as possible specifications for the marketing margin because the (unreported) results of the nested and nonnested tests clearly did not accept these models. Wohlgenant and Mullen reported similar findings for beef, but TL reported that the markup model was preferred to the others.

ing the 1982/83 to 1987/88 seasons.⁴ The monthly data are mostly from November through June for the seasons from 1979/80 to 1987/88.

Increases in the retail price and marketing input prices are expected to widen the marketing margin. Increases in the quantity of navel oranges are expected to widen the marketing margin if the supply curve of marketing inputs is upward sloping, which is plausible in the short run.

A larger proportion of fresh navels shipped during the later months of the marketing season is of lower grades, which results in greater shrink and larger price spreads for these oranges. The monthly binary variables account for the positive influence of these lower quality navels on the average price spread. CA Valencias is a substitute for navels which arrive on the market beginning in April. Its effect is indirectly measured by the influence on the retail price, which would fall when Valencias arrive and, thus, decrease the price spread. The negative influence of Valencias on retail prices is also partially imbedded in the monthly binary variables if their effect on retail prices is different from other demand shift variables.

Data and Estimation Procedures

Monthly U.S. retail prices are used in this analysis instead of weekly retail prices for the cities used by TL. City prices may reflect unique spatial factors. Because the data of this analysis cover a longer time period than TL's study, the data of this analysis display greater variation of market conditions. Such variation reduces the influence of conditions unique to one season on the findings. The monthly retail prices also allow for estimating the influence of a suspension for the United States rather than a few cities. Monthly data are adequate for this type of analysis, although prorate is administered weekly, because the NOAC revises its tentative weekly handler prorates about every month. Actual weekly prorates and tentative prorates are about equal.

⁴ Following the U.S. Department of Agriculture's development of program guidelines for administering marketing order regulations in early 1982, the department and the Navel Orange Administrative Committee (NOAC) for the first time suspended prorates before the end of the season. A suspension occurs when NOAC does not use prorate in the industry. The share of the annual crop shipped during the suspension varied from 12% to 53% during the 1982/83 to 1987/88 seasons. TL artificially distinguishes the 1984/85 suspension from the others by implicitly arguing that the 1984/85 suspension was entirely unexpected because it was announced only days before it went into effect. TL presumably thought that the effect of a "planned" suspension is different from an "unexpected" one. This distinction can be important conceptually, but, in practice, it is groundless for two reasons. First, USDA and industry officials had discussed plans to suspend prorate later in the 1984/85 season. Second, farm prices exceeded parity and were projected to do so for the remainder of the 1984/85 season, and the secretary of agriculture is instructed by the enabling federal marketing order legislation to suspend regulations under such circumstances.

Retail prices for fresh navels in cents per pound are from the Bureau of Labor Statistics (BLS), U.S. Department of Labor (USDOL). These average prices are collected throughout the month by surveying actual prices in supermarkets in eighty-five U.S. cities. These prices are weighted by the estimated stores' sales, which are weighted by navel sizes and grades. The marketing input price is an average of the producer price index for number two diesel fuel (1982-84 = 100), average hourly earnings index for non-supervisory workers in wholesale groceries and related products (1982-84 = 100), and average hourly earnings index for nonsupervisory workers in retail grocery stores (1982-84 = 100). These indexes are from USDOL. Weekly shipments of fresh CA navels to the domestic market (defined by the order as the continental U.S. and Canada) in million pounds are from NOAC's annual bulletin; they were used to construct monthly fresh-use shipments.⁵ Weekly shipments were lagged one week because three to seven days usually are needed to haul navels to retailers from handlers. Such lagging of shipments and FOB prices, which is consistent with Ward's analysis, accounts for transporting and handling navels. As a result the FOB prices for navel shipments more closely match with the retail prices for the same shipments. Monthly FOB prices for fresh CA navels in cents per pound are constructed by weighting the one-week lag of FOB prices by the one-week lag of fresh-use shipments. Weekly FOB prices are reported in weekly NOAC bulletins. These prices are collected from representative handlers and weighted by the handlers' fresh-use shipments, which are weighted by sizes and grades.⁶ Price data (*M*, *RP*, and *MIP*) are deflated by the consumer price index for all items (1982-84 = 1.00).

Because several suspensions began during the month rather than at the beginning or end of the month, a criterion was needed to identify these months as either a suspension or prorate month (SUSP).⁷ The three months that had between four to eleven days with the one-week lag of shipments from a suspension were identified as prorate months because the bulk of the one-week lag of shipments were from a prorate pe-

⁵ Data on shipments to Canada are unavailable, but industry personnel indicate that such shipments are a small share of domestic shipments.

⁶ For FOB price for fresh CA navels, and not the grower level price for navels, was used to compute the marketing margin (even though the enabling legislation for marketing orders is intended to affect farm prices), because the difference between the FOB and grower's prices, as reported by National Agricultural Statistics Service (NASS), USDA, is constant over the months of a marketing season. Consequently, the effect of a prorate suspension on marketing margins is the same, whether the FOB or grower's prices are used to compute the marketing margin.

⁷ Eleven of the 55 observations are from a suspension. TL had 42 observations for each city over a data period covering 77 weeks. TL consequently had about 8 or 9 observations during their 16-week suspension period, if the number of observations during the 1984/85 suspension is in proportion to the 42 observations over the data period covering 77 weeks.

riod. The one month that had seven days with the one-week lag of shipments from a prorate period was identified as a suspension.⁸

In contrast to the approach in TL's analysis, retail and FOB prices and shipments in the marketing margin equation are endogenous variables determined simultaneously with handlers' supply and the retail demand for navels.⁹ Predetermined variables appearing in the retail demand and supply as well as the marketing margin equations are normally used as instrumental variables when estimating the marketing margin equation (Johnston). The supply of fresh-use navels is heavily determined by the prorate volume (Thor and Jesse). When making the prorate decision, growers and handlers consider the price, volume ready for shipment, carryin, and other variables. The quantity of navels demanded at the retail level is affected by the price, income, and other variables. The per capita disposable income, carryin of navels, marketing input prices, and all binary variables consequently are the predetermined variables for this abbreviated model.¹⁰ These predetermined variables were used as the instrumental variables in the two-stage least squares (TSLS) estimation.

The small value of the Durbin-Watson (D-W) statistic in the preliminary analysis suggested a positive first-order autocorrelation scheme in the residuals. This autocorrelation may be partly caused by expanding exports. For models with first-order autoregressive errors, Fair indicates that the instrumental variables should include the predetermined variables in the system and also the one-period lag of all endogenous and predetermined variables. The marketing margin equation then are estimated by TSLS correcting for first-order autocorrelation in the residuals using the Cochrane-Orcutt procedure (Johnston).¹¹ The instrumental variables included the variables suggested by Fair.

Empirical Findings

The estimated results of the price spread model are

$$M_t = 1.436 + 0.461*RP_t + 0.035*RP_tQ_t \\ (0.272) \quad (9.463) \quad (2.611)$$

⁸ There were a few early months in several seasons when NOAC voluntarily did not use prorate because shipments are small and the crop had not matured, but these months did not pose any problem since retail prices were not available.

⁹ The prices for marketing inputs and services are assumed predetermined.

¹⁰ U.S. per capita disposable income in dollars is from the *Survey of Current Business*, U.S. Department of Commerce. It was deflated by the consumer price index for all items. Carryin was constructed by subtracting the actual total shipments prior to the beginning of the current month from the current season's estimated crop. Shipments and intraseasonal crop estimates are from the weekly and annual NOAC crop reports; they were lagged one week.

¹¹ The coefficients for the equations estimated by TSLS correcting for first-order autocorrelation are negligibly different from the coefficients estimated by OLS correcting for first-order autocorrelation. Thus, the findings are not sensitive to alternative estimation procedures.

$$+ 0.511*MIP_t + 0.381*APR_t \\ (0.108) \quad (1.070)$$

$$+ 2.522*MAY_t + 5.393*JUN_t \\ (3.570) \quad (5.045)$$

$$+ 0.972*SUSP_{t-1} \\ (2.108)$$

$$\text{Rho} = 0.798, R^2 = 0.892, \text{D-W} = 2.012. \\ (9.737)$$

The estimated coefficients have the expected signs and most are significantly different from zero based on asymptotic *t*-tests (in parentheses) at the 0.05 significance level. The estimated coefficient for the lagged residual term is denoted by rho, while the coefficient of determination is denoted by R^2 .

In contrast to TL, the estimated coefficient for *SUSP* indicates that the marketing margin increases by nearly 1¢ per pound during a suspension. If monthly estimates of the price elasticities of farm supply and retail demand were available, then the distribution of effects among handlers, retailers, and consumers could be approximated (Gardner). While estimates of such elasticities are not readily available, the magnitude of the margin increase is suggested by comparing this increase to several prices. For example, the 1¢ per pound increase in the marketing margin equals 3.7% of the average marketing margin, 2.3% of the average retail price, and 5.8% of the average FOB price.

In contrast to some of TL's findings, the marketing margin widens when the retail price increases but by much less than the increase in the retail price. Greater volumes also increase the marketing margin. The effect of greater volumes has two opposing effects on the marketing margin. The lower retail price (from greater volumes) reduces the marketing margin, but the greater volumes also elevate the price spread. The net effect on the marketing margin depends on the size of the resulting changes in the retail price and volumes. This analysis did not measure the effect of a suspension on volumes and prices and the resulting impact on price spreads, nor did TL.

Marketing input prices have a statistically insignificant effect on the marketing margin, as in TL's analysis, perhaps because of the relatively low variability displayed by this series.

In contrast to TL, the seasonal binary variables indicate that the price spread increases during the later months of the marketing season because a higher proportion of navels are lower quality, which elevates shrink, and perhaps the impact of competing Valencias.

Concluding Comments

In contrast to TL, this analysis indicates that the average price spread for navels widens during a prorate suspension. While the suspension effect could result

from a shift in market power between retail buyers and handlers, a more plausible explanation is that a greater proportion of the fresh navel shipments are of lower quality when prorate is suspended. The lower quality elevates shrink and, hence, increases marketing costs and the price spread. Since handler prorates tend to restrict shipments to the fresh use (Thor and Jesse, Shepard), handlers have a profit incentive to ship the highest quality navels to the fresh use when prorate is in effect because such navels bring the highest price and shrink the least. When prorates are suspended, profit incentives encourage handlers to ship all navels to fresh use which meet the minimum fresh-use quality requirements and will be sold at a price above the marginal variable marketing cost. Picking and shipping navels which enter processing is profitable for growers even though grower' returns are often negative because navels used in processing represent a part of the production base used in the marketing order from which the maximum quantity of navels eligible for the fresh use is calculated. For a grower, a larger production base means a greater maximum quantity of shipments to the higher-priced fresh use.

[Received January 1990; final revision received June 1990.]

References

- Fair, Ray C. "Efficient Estimation of Simultaneous Equations with Autoregressive Errors by Instrumental Variables." *Rev. Econ. Statist.* 54(1972):444-49.
- Gardner, Bruce L. "The Farm-Retail Price Spread in a Competitive Food Industry." *Amer. J. Agr. Econ.* 57(1975):399-409.
- Johnston, J. *Econometric Methods*, 3rd ed. New York: McGraw-Hill Book Co., 1984.
- Pick, Daniel H., Jeffrey Karrenbrock, and Hoy F. Carman. "Price Asymmetry and Marketing Margin Behavior: An Example for California-Arizona Citrus." *Agribus.* 6(1990):75-84.
- Shepard, Lawrence. "Cartelization of the California-Arizona Orange Industry." *J. Law and Econ.* 29(1986):83-123.
- Thompson, Gary D., and Charles C. Lyon. "Marketing Order Impacts on Farm-Retail Price Spreads: The Suspension of Prorates on California-Arizona Navel Oranges." *Amer. J. Agr. Econ.* 71(1989):647-60.
- Thor, Peter K., and Edward V. Jesse. *Economic Effects of Terminating Federal Marketing Orders for California-Arizona Oranges*. Washington DC: U.S. Department of Agriculture, Econ. Res. Serv. Tech. Bull. No. 1664, Nov. 1981.
- U.S. Department of Agriculture, Agricultural Marketing Service. *Guidelines for Fruit, Vegetable, and Specialty Crop Marketing Orders*. Washington DC, 25 Jan. 1982.
- Ward, Ronald W. "Asymmetry in Retail, Wholesale, and Shipping Point Pricing for Fresh Vegetables." *Amer. J. Agr. Econ.* 64(1982):205-12.
- Wohlgenant, Michael K., and John D. Mullen. "Modeling the Farm-Retail Price Spread for Beef." *West. J. Agr. Econ.* 12(1987):119-25.

Marketing Order Impacts on Farm-Retail Price Spreads: Reply

Gary D. Thompson and Charles C. Lyon

In his comment, Powers presents four main criticisms of our article and attempts to remedy the flaws he perceives by estimating an alternative price spread model. In this reply, we shall first deal with Powers' criticisms and then comment on the adequacy of his revisions on our work and the credibility of his results.

Powers' first criticism is that because of inadequate data, we were unable to capture important quality differences in a "heterogenous" commodity. On the contrary, the weekly FOB price series used in our analysis were constructed by calculating an FOB price weighted by size, grade, and production district price differentials (Thompson and Lyon, p. 651). The three weighted market-level FOB series thus reflected changes in the volume of undesirable sizes and grades. Because FOB prices for only those sizes of oranges for which retail prices were collected were used to construct the FOB series, our weighted FOB prices conformed closely to the prices reported in our retail price series. Thus, the three price series explicitly captured quality effects. Any potential bias introduced by ignoring grade and size changes through time was thereby avoided. As acknowledged by Powers, this bias cannot be remedied with the aggregate data he employs. This is a crucial point because Powers concludes that deteriorating product quality at the end of the season should cause price spreads to increase.

Powers' second criticism suggests that we failed to properly identify weeks during the 1985/86 and 1986/87 marketing seasons when the industry also operated without prorate. Our analysis defined the period of prorate suspension as having occurred when the U.S. Department of Agriculture (USDA) unilaterally suspended prorate during the 1984/85 season. We considered the 1984/85 suspension to be fundamentally different from subsequent periods because of the unique circumstances leading up to the suspension. First, market conditions were markedly different. Freezes in 1984/85 in Florida and Texas left California-Arizona (CA) fresh navel oranges with almost no competing supplies during the 1985 marketing period. Second, a majority of the fresh domestic crop had not yet been shipped. When USDA announced the suspension (beginning the first week in February

of 1985 and extending for 18 consecutive weeks until the end of the season), only 37.6% of the crop had been shipped under prorate (Navel Orange Administrative Committee [NOAC] *Annual Report*, 1985). Third, and most important, the prorate suspension in February 1985 was virtually unanticipated. Industry officials admit that the secretary of agriculture's decision to suspend the prorate caught them off guard.¹

In contrast, removal of prorate during the 1985/86 and 1986/87 seasons was fully anticipated, was of shorter duration, and occurred under much different supply conditions. At the beginning of the 1985/86 season, USDA notified NOAC on 25 November 1985 that USDA would issue weekly volume regulations based upon criteria established by NOAC. After protracted discussions among committee members, NOAC finally agreed on 23 February 1986 to three criteria upon which to base the decision to suspend the weekly prorate.² The prorate in the 1985/86 season was suspended after the week ending on 13 April when two of the three criteria were met. This suspension was clearly anticipated and nearly two-thirds of the fresh domestic shipments had already been accounted for. Rules regarding suspension of the prorate were even more stringent during the 1986/87 season as NOAC proposed on 30 October 1986 that volume controls be imposed in each week through the week ending on 23 April 1987; these proposed regulations were accepted by USDA, and 79.9% of the season's fresh domestic crop was shipped before prorate was suspended. In both these seasons, supply conditions permitted larger shipments from Texas and Florida than during the 1984/85 season. Thus the "suspensions" occurring in the 1985/86 and 1986/87 were clearly different from the 1984/85 suspension in important ways. Acknowledging these dissimilarities, we isolated the 1984/85 suspension as the cynosure of our analysis, thereby permitting examination of industry and market reaction to this unanticipated supply-side policy shock.

¹ Industry officials indicated that the decision to suspend the prorate was not expected because it was one of the more severe options available to the secretary of agriculture and there appeared to be little precedent for such an action.

² NOAC agreed that prorate should be suspended when any two of the following criteria were met: (a) the remaining volume in the largest district—District 1 in central California—was less than 11,500 carloads (26% of the actual shipments from District 1 in 1985/86 or 24% of total shipments from all four districts); (b) weekly fresh domestic shipments of California-Arizona Valencia oranges exceeded 200 carloads; and (c) FOB prices had not declined in three consecutive weeks following the weeks ending on 27 Feb. 1986.

The authors are, respectively, an assistant professor, Department of Agricultural Economics, University of Arizona, and a graduate research assistant, Department of Agricultural and Applied Economics, University of Minnesota.

Powers' third criticism actually embodies two separate but perhaps related issues; dynamic price transmission and simultaneity. In either case, how our results are suspect in light of the Pick, Karrenbrock, and Carman (PKC) study is not made clear. The thrust of the PKC work appears to be in explaining price transmission, not how prorate affect marketing margins. Consequently, the models employed by PKC explain marketing margins only as a linear function of current and past prices; no variables—other than a time trend—purporting to explain margins are included in their specification. Furthermore, in the margin equation specified by PKC the dependent variable is not the FOB-retail margin but rather the change in the margin from one week to the next (Pick, Karrenbrock, and Carman, p. 77). Thus, while our analysis concentrates on explaining the magnitude of the spread between FOB and retail prices, the PKC analysis focuses on changes in this spread. This discrepancy leads us to question strongly the relevance of the PKC results either as a basis to refute our results or as support for the model proposed by Powers. Moreover, if simultaneity is a problem in price spread models as Powers suggests, then the PKC results must also be questioned. Even if the PKC results were directly applicable, however, the results of our diagnostic tests for first-order autocorrelation within and across equations provide indirect evidence that the omission of lagged values of retail prices did not lead to misspecification.

The issue of simultaneity was addressed in our article (Thompson and Lyon, p. 649). We explicitly noted that use of weekly price data precluded specification of a structural model. While aggregation of weekly data into monthly observations would have permitted such a specification, this was undesirable for both practical and theoretical reasons. Practically, aggregation would have resulted in the loss of sufficient data points during the prorate suspension on which to base the analysis. Indeed, the unique nature of the dataset would have been forfeited, rendering a valid "with and without" comparison of industry behavior under marketing order quantity controls infeasible. Because this was the prime advantage of the weekly data, the disaggregated nature of the dataset was retained and the reduced-form equations specified. This enabled us to identify fifteen data points during the 1984/85 suspension period. Monthly data would have reduced the number of data points to four. Furthermore, use of monthly data requires specification of a monthly dummy variable (such as that employed by Powers) tracking suspensions which occur at any week during a given month. This necessarily provides an imprecise measure of the impact of suspensions because both suspension and nonsuspension weekly observations are aggregated into monthly observations.

More important from a theoretical perspective, monthly observations are inconsistent with the institutional setting of the industry. This is because the CA naval orange prorate is a weekly policy instru-

ment used to regulate market flows. Powers' claim that prorates are set largely on a monthly basis is not consistent with the fact that actual prorate allotments are announced only three to ten days prior to the period of regulation (USDA 1988). In addition, announced prorate quantities are commonly revised by NOAC once, twice, and even three times before the end of the weekly period of regulation (NOAC Weekly Bulletins). Clearly then, producers, handlers, and distributors cannot consider tentative prorates to be equivalent to actual prorates. Indeed, if one assumes that weekly prorates and tentative prorates are "about equal," then the unanticipated events of the 1984/85 seasons are not subject to analysis.

To remedy the simultaneity, Powers argues that an instrumental variable estimator is necessary. Because nearly all the price spread models referred to in the literature have been reduced-form models (e.g., Wohlgenant and Mullen), the question of appropriate estimators is seldom addressed. The justification for utilizing an instrumental variable estimator lies in the possible correlation between the explanatory variable and the error term, whether the cause of the correlation derives from simultaneity or other sources. Recent small sample results suggest that instrumental variable estimates may provide little improvement over OLS estimates (Nelson and Startz). In any case, simultaneity would appear to be less of a problem in a short-run model based on weekly observations than in a longer-run model utilizing monthly data. We suspect that retail demand schedules for navel oranges did not shift about radically from week to week. On the supply side, seasonal supply schedules are predetermined for biological reasons, although intra-seasonal supply schedules might conceivably shift. Regardless, these hypotheses can be checked indirectly by comparing the vector of parameter estimates obtained from OLS and two-stage least squares using the Hausman test for errors in variables in single equation models (Hausman).³ We could think of applying an analogous test for the difference between the seemingly unrelated regression estimates and three-stage least squares estimates, but, for the same reasons we could not specify a structural model, we have no appropriate set of instruments with weekly observations for three-stage estimation.

Parenthetically, it should be mentioned that contrary to Powers' assertion, neither the retail price data nor the FOB price data used by PKC were the same as the data used in our analysis. We augmented our retail price series with independently gathered advertised prices. In addition, we employed a different FOB price series than did PKC and Powers. Powers and PKC used an FOB price series reported by NOAC. As noted in our article, we used the FOB prices re-

³ Powers comments that the OLS and two-stage least squares (TSLS) estimates for his model are negligibly different. This result gives circumstantial evidence that the purported impact of simultaneity is itself negligible or that the instruments used in the TSLS estimation did little to purge the endogeneity.

ported in the Western Citrus Report (WCR), which publishes weekly sales prices for every size and grade of navel orange sold by production district. As mentioned above, these data enabled us to adjust for quality differences. Further motivation for the use of the WCR data in lieu of the NOAC data is found in an ERS study (USDA 1985) which notes that the weekly NOAC FOB series represents prices quoted by handlers rather than actual sales prices. This same study states that the NOAC quoted prices are generally viewed within the industry as inaccurate reflections of actual selling prices.

Powers' final criticism that our "explanation" of the influence of prorate suspension on price spreads is "vague" appears to be misplaced. We referred to the difference in opinions between some producers and consumer advocates in the introductory description of the CA navel orange industry because this lack of consensus provided one justification for studying margin behavior under prorate suspension. These conflicting views of price spread behavior were never offered as an explanation in subsequent sections of the article.

Turning now to the alternative specification presented by Powers, the estimation results appear suspect for a number of reasons. The monthly data employed are not only theoretically inconsistent but are, by their very nature, inadequate to capture changes in FOB-retail price spreads. While weekly FOB prices are relatively stable, in-store retail prices fluctuate

significantly from week to week. In addition, supermarkets advertise sale prices periodically for navel oranges as loss leaders. Monthly average retail prices from the Bureau of Labor Statistics do not adequately reflect important weekly changes in navel orange prices. Aggregation across eighty-five cities suggests as well that potentially important spatial differences cannot be accounted for. Failure to capture the true extent of price variability, particularly that variability resulting from quality changes, may be one reason for the increasing margins at season's end indicated in Powers' estimation results.

Reestimation of the Markup Model

Of Powers' four criticisms, the "failure" to properly identify suspension periods is potentially the most damaging. We have already indicated why the 1984/85 suspension was unique. We were curious, nonetheless, to explore how our estimates might have changed had we included all other "suspensions," as Powers exhorts. Accordingly, we reestimated two variations on our original markup model. In the first, we included a dichotomous variable, *DSUSP*, set equal to one for any week in which prorate was not in effect during all three seasons. In the second, we decomposed the effects of removing the prorate by including two dummy variables, one for weeks during the 1984/85 suspension (*D1984/85*) and a second

Table 1. Maximum Likelihood Estimates of the Markup Model with Various Dummies

Dependent Variable	Independent Variables							
	Constant	<i>P</i>	<i>TR</i>	<i>D1984/85</i>	<i>DOTHER^a</i>	<i>DSUSP</i>	<i>DMAY</i>	<i>TIME</i>
Original Markup Model								
<i>MAT</i>	-0.03 (-1.14) ^b	0.99 (11.96)	-5.52E-5 (-1.58)	-0.013 (-2.02)			-0.024 (-4.47)	0.0001 (1.39)
<i>MDL</i>	0.072 (1.19)	0.55 (4.31)	-1.62E-5 (-1.37)	0.008 (0.81)			-0.002 (-2.14)	0.0004 (2.39)
<i>MSF</i>	-0.178 (-1.75)	1.05 (12.94)	0.0006 (1.01)	-0.012 (-1.74)			-0.031 (-4.78)	0.0003 (1.96)
Single Suspension Dummy (<i>DSUSP</i>):								
<i>MAT</i>	-0.04 (-1.48)	0.91 (11.97)	-0.000047 (-1.39)			-0.00002 (-0.00)	-0.028 (-4.89)	0.0003 (4.16)
<i>MDL</i>	0.099 (1.59)	0.59 (4.45)	-2.15E-4 (-1.74)			0.005 (0.63)	-0.023 (-2.07)	0.0004 (2.74)
<i>MSF</i>	-0.024 (-2.15)	0.91 (11.11)	9.93E-4 (1.45)			0.004 (0.85)	-0.032 (-4.76)	0.0005 (3.89)
Multiple Suspension Dummies (<i>D1984/85</i> and <i>DOTHER</i>):								
<i>MAT</i>	-0.05 (-2.29)	1.25 (14.66)	-4.00E-5 (-1.57)	-0.028 (-5.04)	0.01 (2.46)		-0.022 (-4.91)	-0.00009 (-1.11)
<i>MDL</i>	0.02 (0.28)	0.64 (5.31)	-5.00E-5 (-0.48)	0.00064 (0.06)	0.0013 (0.16)		-0.03 (-2.69)	0.0003 (1.86)
<i>MSF</i>	-0.20 (-1.65)	1.01 (11.47)	8.00E-4 (1.13)	-0.014 (-1.90)	0.0099 (1.78)		-0.03 (-4.01)	0.0002 (0.16)

^a *DOTHER* equals one when prorate is suspended in the 1985/86 and 1986/87 seasons; otherwise *DOTHER* equals zero.

^b *t*-statistics are in parentheses.

for non-prorated weeks occurring in the 1985/86 and 1986/87 seasons (*DOTHER*). All remaining data and variables are identical to those referred to in our article.

Had we mistakenly lumped all prorate "suspensions" into the same dichotomous variable (*DSUSP*), the effects of removing prorate would have appeared to be insignificant for all three cities (see table 1). The magnitude of the estimates and their large standard errors do not provide much basis for inferring the validity of Powers' results or ours. In fact, had we initially obtained these results, we would likely not have attempted to publish them. Alternatively, when two separate dummies for non-prorated periods in the 1984/85 (*D1984/85*) and 1985/86–1986/98 (*DOTHER*) seasons are specified, decreases in the price spreads for Atlanta and San Francisco in the 1984/85 season are strikingly similar to those reported in our article. These results indicate that the effect of the 1984/85 prorate suspension was clearly different from non-prorated periods in subsequent seasons. These empirical results are consistent with the fact that the circumstances surrounding the 1984/85 prorate suspension were substantially different from other prorate "suspensions."

We might speculate that with data from additional seasons included in the empirical analysis, the positive impact on price spreads of additional "suspensions" which occur at the end of each season would ultimately swamp the negative impact of the 1984/85 suspension if all suspension periods are included in single dummy variable. The estimation results of Powers' price spread model tend to substantiate this speculation. Yet, in light of the unique nature of the 1984/85 suspension, we continue to conclude that FOB–retail price spreads in two geographically disparate cities declined significantly when the prorate was suspended in 1985.

Finally, we would like to take this opportunity to correct a typographical error in the original article. In table 2, the first independent variable in the relative model should be price.

[Received August 1990; no revision.]

References

- Hausman, J. A. "Specification Tests in Econometrics." *Econometrica* 46(1978):1251–71.
- Navel Orange Administrative Committee (NOAC). *Annual Reports*, Los Angeles CA, 15 Nov. 1985, 5 Jan. 1987, and 14 Aug. 1987.
- Nelson, Charles R., and Richard Startz. "Some Further Results on the Exact Small Sample Properties of the Instrumental Variable Estimator." *Econometrica* 58(1990):967–76.
- Pick, Daniel H., Jeffrey Karrenbrock, and Hoy F. Carman. "Price Asymmetry and Marketing Margin Behavior: An Example for California-Airzona Citrus." *Agribus* 6(1990):75–84.
- Thompson, Gary D., and Charles C. Lyon. "Marketing Order Impacts on Farm-Retail Price Spreads: The Suspension of Prorates on California-Arizona Navel Oranges." *Amer. J. Agr. Econ.* 71(1989):647–60.
- U.S. Department of Agriculture. *Before the Secretary of Agriculture, Proceedings*. AMA Docket No. F&V 907-6 to 10, Washington DC, 29 Jan. 1988.
- U.S. Department of Agriculture, Economic Research Service. "Economic Implications of the 1984/85 California-Arizona Navel Orange Shipment Prorate Suspension." Mimeographed. Washington DC, 15 Sep. 1985.
- Wohlgenant, Michael K., and John D. Mullen. "Modeling the Farm-Retail Price Spread for Beef." *West. J. Agr. Econ.* 12(1987):199–225.

Rice in Asia: Is It Becoming an Inferior Good?

Comment

Jikun Huang, Cristina C. David, and Bart Duff

In a recent article in this *Journal*, Ito, Peterson, and Grant (IPG) investigated the nature of demand for rice in fourteen Asian countries using a generalized covariance demand model. This model was estimated from pooled time-series-cross-section data from 1961 to 1985 for three groups of countries classified by changes in per capita rice consumption levels. Rice consumption per capita was specified as a function of per capita gross domestic product (GDP) in real terms and the ratio of world price of rice to wheat. Because income elasticities tend to be lower at higher income levels, a log-inverse-log model was adopted and ridge regression was employed to minimize multicollinearity problems between $\ln Y$ and Y^{-1} . The results indicated that income elasticities have significantly declined. Moreover, income elasticities are negative in half of the countries studied, leading to the conclusion that rice is becoming an inferior good in Asia.

The fact that income elasticities for rice and other staple food decline as income increases has long been established (Alderman, Bouis 1989). However, for comparable years, IPG's estimates are much lower than country-specific estimates of income elasticities in Asia based on either cross-section or time-series data, which are nearly all positive (table 1). While the latter pertain mostly to estimates for the 1970s, negative income elasticities for low-income countries such as Nepal and Bangladesh seems implausible. Even for a rapidly developing country such as Thailand, income elasticities may be expected to become negative only sometime in the 1980s. Only for Japan, Taiwan, Singapore, and Malaysia, where per capita incomes are relatively high, can negative income elasticities be expected. This comment demonstrates that problems with definition of income and price variables as well as application of ridge regression have led to underestimation of income elasticities in the IPG study.

Problems of Estimation

Estimation problems include such specification issues as those related to income and price and the sometimes misleading results of ridge regression.

The authors are, respectively, a research fellow, and agricultural economists, the International Rice Research Institute, Manila, Philippines.

The authors are grateful to Yujiro Hayami, Keijiro Otsuka, and the *Journal* reviewers for their suggestions.

Specification Issues

Income. IPG defined income as per capita GDP in real terms measured in domestic currencies. With cross-country pooled data, this means that rice consumption was regressed on income measured in different currencies. In a generalized covariance model, OLS estimate of the parameters, $\beta = [X'X]^{-1}X'Y$, is invariant to differences in measurement units across countries. However, in ridge regression, β is estimated as $[X'X + kI]^{-1}X'Y$, where k is defined as the ridge factor. Addition of the constant value k to all elements in the $X'X$ matrix measured in different currencies will result in a different estimate of the parameters for the pooled cross-country data from separate regressions for each country. Differences in estimated income elasticities across countries therefore will reflect not only the effect of preference but also differences in the unit of measurement of income.

Another problem is use of per capita gross domestic product (GDP) rather than disposable income or private consumption expenditure which better reflect the budget constraint of consumers. GDP includes government expenditures which usually increases faster than other components of domestic product. Hence, use of GDP instead of more appropriate measures of consumers' income will result in underestimation of income elasticity.

Price. Consumer demand for a commodity depends on, among others, domestic prices. In most Asian countries, the government intervenes extensively in the rice market and, to a lesser extent, also in the wheat market. Thus, domestic prices of these commodities will diverge from world market prices in varying degrees across countries and over time. Indeed, variations in the price ratio of rice to wheat across countries as a result of differential rates of nominal protection between the two commodities ranged from 0.20 in Burma to 3.84 in South Korea in 1986. Conversely, the ratio of world prices increased by 31% from 1961 to 1985, whereas in twelve out of the fourteen countries, the ratio of domestic prices increased ranging from 3% in Nepal to 178% in South Korea in the same year. Only Singapore and Burma had declining ratios of domestic prices. IPG's use of world prices instead of domestic prices is inappropriate in such policy environment.

Table 1. Income Elasticity Estimates for Rice in Asia

Country	Author	Income Elasticity	Expenditure Survey Year	IPG ^a
Bangladesh	Bouis	0.34–0.83 ^b	1973	–0.04
Burma	FAO	0.10	1971	0.03
China	FAO	0.40	1971	0.25
India	Bouis	0.05–0.85 ^b	1977	0.14
	ADB	0.77	1985	0.13 ^a
Indonesia	FAO	0.70	1971	0.26
	Timmer & Alderman	0.28–1.16 ^b	1976	0.19
	Bouis	0.04–0.40 ^b	1981	0.12
Japan	FAO	–0.10	1971	–0.56
Korea	FAO	0.30	1971	0.06
Malaysia	FAO	0.20	1971	–0.09
	ADB	–0.10	1974–85	–0.50
Nepal	FAO	0.30	1971	–0.37
Philippines	FAO	0.20	1971	0.15
	Quisumbing	0.55–1.71 ^b	1978	0.11
Singapore	FAO	0.10	1973	–0.33
Thailand	Trairatvorakul	0.0–0.41 ^b	1975	–0.25
	Bouis	–0.14–0.20 ^b	1975	–0.25
Taiwan	FAO	0.30	1971	–0.39

Sources: Asian Development Bank, IPG (tables 3 and 6); Quisumbing, Bouis (1990, tables 28 and 75).

^a Refers to estimates by IPG pertaining to the same expenditure survey year except for the second line for India which refers to income elasticity in 1984.

^b The range relates to income level (high to low). Elasticity for the Philippines estimated by Quisumbing is food expenditure elasticity of demand for rice.

Table 2. Selected Estimates and Statistics in Ridge Regression

	Inverse of Income		Log of Income		Log of Price	
	Base	Dummies	Base	Dummies	Base	Dummies
Group I						
Taiwan	–242.6 (126.9) ^a		–0.691 (.115)		0.019 (.165)	
Japan		–525.2 (335.4)		0.031 (.017)		0.022 (.257)
Malaysia		–298.9 (140.0)		0.024 (.014)		–0.336 (.295)
Nepal		–190.3 (145.6)		0.002 (.035)		–0.038 (.590)
Singapore		–346.2 (214.4)		0.013 (.011)		–0.445 (.213)
Thailand		–209.4 (97.3)		0.034 (.012)		–0.154 (.215)
$k = .002$	$R^2 = .83$	$F = 26.4$	$n = 152$	$D.F. = 128$		
Group II						
India	–38.7 (31.1)		0.147 (.056)		–0.168 (.065)	
Bangladesh		74.9 (23.3)		0.057 (.006)		–0.069 (.090)
S. Korea		49.5 (31.2)		0.032 (.006)		0.022 (.065)
$k = .01$	$R\text{-square} = .96$	$F = 179.1$	$n = 84$	$D.F. = 72$		
Group III						
Burma	–53.3 (86.7)		0.342 (.184)		–0.021 (.030)	

Table 2. Continued

	Inverse of Income		Log of Income		Log of Price	
	Base	Dummies	Base	Dummies	Base	Dummies
China		35.6 (56.5)		-0.065 (.021)		0.519 (.358)
Indonesia		-47.6 (65.7)		-0.029 (.020)		-0.009 (.103)
Philippines		-8.9 (205.5)		-0.081 (.025)		-0.141 (.120)
$k = .0005$	$R\text{-square} = .91$	$F = 60.9$	$n = 107$	$D.F. = 91$		

Note: Standard t -test is not relevant in ridge regression.

* Standard errors appear in parentheses.

Ridge Regression

Brown and Beattie have shown that ridge estimation of some models can give poor or misleading results. Significantly biased coefficients will be estimated when the model includes: (a) positively correlated variables that have true coefficients with opposite signs and highly unequal magnitudes or (b) negatively correlated variables that have true coefficients with the same sign but large unequal magnitudes.

Although all variables may not be positively or negatively correlated in a generalized log-inverse-log model with country dummy variables, the bias can be very large due to the special nature of the model. Besides price and income, four other sets of variables are highly intercorrelated, i.e., $\{y^{-1}, \ln y\}$, $\{d_i y^{-1}, d_i\}$, $\{d_i \ln y, d_i\}$, and $\{d_i y^{-1}, d_i \ln y\}$, where d_i is the dummy variable for country i . In addition, the magnitudes of the coefficients for these correlated variables are very different when income elasticities (e) are assumed to change over time.¹ Use of ridge regression in the IPG study without incorporating a priori information may, therefore, have led to biased estimates.

A Reestimation

Because IPG's analysis is subject to a number of econometric problems and the extent to which measurement problems of income and prices affect the estimates of income elasticities can only be determined empirically, the model was reestimated with several changes. First, income is defined as per capita private consumption expenditure (PCE) in real terms (deflated by consumer price index) and expressed in U.S. dollars converted by the purchasing power parity (PPP) exchange rate for 1980.² Second, the ratio of domestic price of rice to wheat flour has been used. Third, Sri Lanka is excluded from the

analysis because of limitation of price data. Data for rice consumption, population, and income are from the same sources as IPG's except for the PCE for Bangladesh, which is from the Asian Development Bank. Domestic prices of rice and wheat are from individual country sources. The PPP's are from Pardey and Roseboom.

To minimize possible bias in use of ridge regression in the log-inverse-log demand model, the following additional a priori information were considered in choosing the optimal k -value: (a) k should be small;³ (b) the coefficient of Y^{-1} is expected to be negative;⁴ and (c) average level of estimated income elasticities (e) obtained from log-inverse-log model over the study period should not be very far from the constant elasticities (e^0) derived from the double-log demand model using OLS.⁵

The general procedure for choosing the optimal k -value are as follows. Based on k values ranging from 0 to 1, a set of k 's, say A , which generate relatively stable trace is identified. From this, a subset B , is selected such that the condition of negative coefficient of Y^{-1} is generally satisfied. Finally, the "optimal" k is chosen to be the lowest k which provide estimates that largely satisfy condition (c).

Results

The estimates of coefficients based on the revised measures of income and price as well as alternative choice of k values are presented in table 2, and the estimated income elasticities are reported in table 3. The new estimates of income elasticities are, in general, higher than the IPG results. Only Japan, Malaysia, Singapore, Taiwan, and Thailand have neg-

¹ Because $e = c - b/Y$, where c is the coefficient of $\ln Y$ and b of Y^{-1} , then b is expected to have a much large value than c . Suppose $e_{1961} = .3$, $e_{1987} = .1$, $Y_{1961} = 1000$ and $Y_{1987} = 2000$; then $b = -400$, $c = -.1$ and $b/c = 4000$.

² In China, national income is used due to lack of reliable data.

³ Because it is undesirable to introduce excessive bias, one should be conservative in increasing k , particularly in the case of the log-inverse-log dummy demand model as explained above.

⁴ This is based on the earlier observation that income elasticities decline as income increases.

⁵ While the double-log model does not allow for changing income elasticities, it provides estimates derived by unbiased OLS estimation which can serve as a reference for the average income elasticity.

Table 3. Income Elasticities for Rice in Asian Countries, 1961-88

Year	Group I					Group II			Group III				
	Taiwan	Japan	Malaysia	Nepal	Singapore	Thailand	India	Bangladesh	Korea	Burma	China	Indonesia	Philippines
1961	-0.237	-0.227	-0.047		-0.284	0.191	0.254	0.129	0.150	0.495	0.370		0.344
1962	-0.253	-0.258	-0.064		-0.284	0.164	0.255	0.124	0.153	0.494	0.379		0.342
1963	-0.261	-0.288	-0.067		-0.288	0.132	0.249	0.126	0.155	0.461	0.371		0.339
1964	-0.306	-0.324	-0.089		-0.257	0.112	0.246	0.122	0.158	0.473	0.360		0.339
1965	-0.334	-0.334	-0.103		-0.264	0.071	0.252	0.120	0.157	0.487	0.350		0.337
1966	-0.344	-0.366	-0.113		-0.282	0.029	0.248	0.116	0.158	0.515	0.341	0.663	0.336
1967	-0.367	-0.389	-0.106		-0.302	-0.003	0.244	0.120	0.160	0.514	0.348	0.596	0.334
1968	-0.383	-0.408	-0.115		-0.334	-0.019	0.249	0.114	0.161	0.490	0.355	0.591	0.332
1969	-0.387	-0.427	-0.142		-0.356	-0.025	0.245	0.119	0.162	0.482	0.344	0.577	0.329
1970	-0.404	-0.440	-0.157		-0.390	-0.110	0.248	0.121	0.164	0.475	0.333	0.581	0.331
1971	-0.421	-0.449	-0.162		-0.420	-0.099	0.246	0.130	0.165	0.477	0.331	0.584	0.334
1972	-0.445	-0.466	-0.176		-0.438	-0.120	0.244	0.090	0.166	0.480	0.330	0.560	0.334
1973	-0.475	-0.478	-0.200		-0.444	-0.153	0.243	0.082	0.168	0.486	0.327	0.548	0.333
1974	-0.436	-0.472	-0.211		-0.436	-0.152	0.250	0.093	0.169	0.478	0.328	0.537	0.332
1975	-0.478	-0.477	-0.193	0.268	-0.441	-0.166	0.256	0.126	0.170	0.491	0.325	0.540	0.330
1976	-0.487	-0.481	-0.219	0.230	-0.455	-0.191	0.247	0.111	0.171	0.502	0.327	0.537	0.329
1977	-0.499	-0.486	-0.247	0.373	-0.461	-0.208	0.243	0.093	0.171	0.496	0.324	0.527	0.326
1978	-0.513	-0.493	-0.279	0.324	-0.469	-0.228	0.240	0.118	0.172	0.483	0.319	0.526	0.323
1979	-0.527	-0.502	-0.301	0.260	-0.479	-0.240	0.240	0.118	0.173	0.486	0.317	0.496	0.321
1980	-0.535	-0.502	-0.335	0.340	-0.488	-0.239	0.226	0.115	0.173	0.477	0.315	0.470	0.321
1981	-0.537	-0.502	-0.338	0.361	-0.490	-0.235	0.227	0.129	0.173	0.470	0.314	0.466	0.321
1982	-0.542	-0.506	-0.338	0.367	-0.493	-0.236	0.227	0.060	0.173	0.464	0.312	0.459	0.321
1983	-0.549	-0.511	-0.345	0.446	-0.499	-0.254	0.226	0.116	0.173	0.465	0.309	0.455	0.320
1984	-0.558	-0.514	-0.355	0.335	-0.507	-0.267	0.226	0.120	0.173	0.460	0.306	0.452	0.322
1985	-0.564	-0.517	-0.350	0.356	-0.503	-0.267	0.225	0.123	0.174	0.463	0.303	0.448	0.327
1986	-0.570	-0.520	-0.305	0.413	-0.509	-0.281	0.223	0.121	0.174	0.468	0.301	0.450	0.328
1987	-0.580	-0.525	-0.306	0.435	-0.522	-0.302	0.231	0.121	0.174	0.498	0.299	0.446	0.325
1988	-0.591	-0.530	-0.349			-0.328	0.237	0.125	0.174	0.524	0.299	0.446	0.324

Table 4. Sensitivity of Estimates of Income Variables in Ridge Regression

Country	Variable	$k = 0$	$k = .00001$	$k = .002$	$k = .1$	$k = .5$
Taiwan	$y - 1$	-393	-571	-243	86	58
Japan	$d1y - 1$	-1999	-1656	-525	305	151
Malaysia	$d2y - 1$	-2631	-820	-299	-3	13
Nepal	$d3y - 1$	-687	-114	-190	-73	-41
Singapore	$d4y - 1$	-5201	-2419	-346	78	-44
Thailand	$d5y - 1$	-343	-171	-209	-21	43
Taiwan	$\ln y$	-0.88	-1.05	-0.69	-0.17	-0.09
Japan	$d1\ln y$	-0.27	-0.04	0.03	-0.00	-0.00
Malaysia	$d2\ln y$	-1.75	-0.28	0.02	0.00	-0.00
Nepal	$d3\ln y$	-1.11	-0.01	0.00	-0.02	-0.02
Singapore	$d4\ln y$	-2.22	-0.79	0.01	-0.01	-0.01
Thailand	$d5\ln y$	-0.12	0.04	0.03	0.02	0.01

	Variable	$k = 0$	$k = .00001$	$k = .01$	$k = .1$	$k = .5$
India	$y - 1$	1271	137	-39	-64	-44
Bangladesh	$d1y - 1$	-986	29	75	85	65
Korea (S)	$d2y - 1$	-1247	-105	49	82	71
India	$\ln y$	3.04	0.39	0.15	0.09	0.06
Bangladesh	$d1\ln y$	-2.11	0.21	0.06	0.04	0.03
Korea (S)	$d2\ln y$	-2.80	-0.13	0.03	0.02	0.01

	Variable	$k = 0$	$k = .00001$	$k = .0005$	$k = .1$	$k = .7$
Burma	$y - 1$	2987	102	-53	-41	-20
China	$d1y - 1$	-3010	-113	36	-53	-45
Indonesia	$d2y - 1$	-3171	-100	-48	-103	-40
Philippines	$d3y - 1$	-2903	229	-9	-220	-132
Burma	$\ln y$	8.66	0.80	0.34	0.07	0.01
China	$d1\ln y$	-8.40	-0.50	-0.06	-0.03	-0.02
Indonesia	$d2\ln y$	-8.51	-0.28	-0.03	-0.00	-0.00
Philippines	$d3\ln y$	-8.26	-0.12	-0.08	-0.03	-0.02

Table 5. Sensitivity of Income Elasticities in Ridge Regression

Country	Year	Elasticities				
		Constant	$k = 0$	$k = .00001$	$k = k^{**}$	$k = .1$
Taiwan	1961	-0.526	-0.144	0.018	-0.237	-0.332
	1975	-0.526	-0.534	-0.549	-0.478	-0.247
	1988	-0.526	-0.718	-0.816	-0.591	-0.207
Japan	1961	-0.431	0.207	0.165	-0.227	-0.393
	1975	-0.431	-0.575	-0.563	-0.477	-0.265
	1988	-0.431	-0.738	-0.715	-0.530	-0.239
Malaysia	1961	-0.230	0.831	0.264	0.047	-0.266
	1975	-0.230	0.016	-0.111	-0.193	-0.244
	1988	-0.230	-0.860	-0.514	-0.349	-0.220
Nepal	1975	0.553	0.399	0.455	0.268	-0.224
	1980	0.553	0.577	0.569	0.340	-0.226
	1987	0.553	0.814	0.719	0.435	-0.228
Singapore	1961	-0.412	-6.841	0.162	-0.284	-0.295
	1975	-0.412	-5.351	-0.635	-0.441	-0.251
	1987	-0.412	-4.585	-1.044	-0.522	-0.229
Thailand	1961	-0.096	-2.380	0.386	0.191	-0.275
	1975	-0.096	-1.798	-0.202	-0.166	-0.224
	1988	-0.096	-1.535	-0.467	-0.328	-0.201
India	1961	0.071	-0.482	0.007	0.254	0.270
	1975	0.071	-0.556	-0.001	0.256	0.274
	1988	-0.071	0.090	0.069	0.237	0.241

Table 5. Continued

Country	Year	Elasticities				
		Constant	$k = 0$	$k = .0001$	$k = k^*$	$k = .1$
Bangladesh	1961	0.137	0.338	0.252	0.129	0.090
	1975	0.137	0.322	0.243	0.126	0.089
	1988	0.137	0.307	0.234	0.125	0.088
Korea	1961	0.218	0.283	0.173	0.150	0.065
	1975	0.218	0.327	0.231	0.170	0.097
	1988	0.218	0.336	0.245	0.174	0.105
Burma	1961	0.554	0.116	0.307	0.495	0.184
	1975	0.554	0.324	0.514	0.491	0.181
	1988	0.554	-1.512	0.452	0.524	0.206
China	1961	0.321	0.382	0.354	0.370	0.530
	1975	0.321	0.324	0.324	0.325	0.291
	1988	0.321	0.291	0.308	0.299	0.156
Indonesia	1966	0.517	0.781	0.516	0.663	0.562
	1977	0.517	0.534	0.518	0.527	0.369
	1988	0.517	0.387	0.519	0.446	0.253
Philippines	1961	0.302	0.285	0.236	0.344	0.379
	1975	0.302	0.304	0.312	0.330	0.319
	1988	0.302	0.313	0.344	0.324	0.293

¹ $k^* = .002$, $.01$, and $.0005$ for group 1, group 2, and group 3, respectively.

ative income elasticities among the thirteen countries analyzed.⁶ Income elasticities for Bangladesh and Nepal are now positive. For Nepal, elasticities have increased over time because per capita consumption expenditure in real terms apparently declined. Korea shows a slightly rising income elasticity which may be explained by the significant substitution of rice for barley in the diet during the study period which has not been taken into account in this simple model.⁷ Finally, among countries with positive income elasticities, these have declined more gradually than the IPG estimates.

Sensitivity Analysis

The estimated coefficients of income variables and thus, the derived income elasticities are sensitive to the choice of k -value (tables 4 and 5). Based on the ridge trace rule, optimal k -value might be 0.1 for all three country groups. However, at $k = 0.1$, some prior information is violated in many countries. The coefficients of Y^{-1} for all countries in group I are positive implying that income elasticities increase as incomes rise. Furthermore, average estimated income elasticities over the study period for Nepal and Thailand are far from those derived using double-log model. In group II and group III, most countries' e are far from e^0 . By incorporating three prior infor-

mation sets and following the estimation procedures outlined earlier, we chose $k = .002$ as optimal for group I. In group II, since only positive coefficients of Y^{-1} can be obtained for Bangladesh and Korea over any k ranging from zero to one, $k = .01$ which provides the relative small positive value coefficients and meet all other criteria is chosen. At $k = .0005$, all criteria are satisfied for group III countries.

Concluding Remarks

Our analysis indicates the IPG estimates of income elasticity are significantly understated.⁸ While income elasticities will undoubtedly decline as income increases, only Japan, Taiwan, Singapore, Malaysia, and Thailand have income levels that support negative estimates of income elasticities for rice. However, the population and rice consumption of these five countries account for less than 10% of the totals for Asia. In most Asian countries, therefore, rice is not an inferior good; and income elasticities for rice, given the relatively stable elasticity trend in table 3, will likely remain positive throughout the 1990s.

The implication drawn by IPG that "there is a potential for excess supplies of rice to develop in Asia, putting downward pressure on rice price" is not only

⁶ The negative income elasticities for Thailand in the 1960s and 1970s remain implausible. The underreporting of rice exports due to high export taxes during that time may have understated per capita rice consumption in the early period.

⁷ This substitution has run its course, and thus rice consumption per capita may be expected to decline in Korea in the future.

⁸ Bouis (1990), in another comment on the IPG's study, argued that growing urbanization and increasing commercialization of rice production, not simply increasing income per se can explain decreasing national per capita rice consumption. If these factors are properly taken into account, revised estimates of income elasticities are expected to be even higher. This is confirmed by our preliminary estimates of an Asian cereal grains demand model that includes urbanization effect.

based on faulty estimation but a very simplistic view of the rice situation. Future rice supply prospects are not expected to be strong (Levine et al., David). The land frontier in most Asian countries is essentially closed. Crop area planted to rice has not grown throughout the 1980s. The easy gains from the green revolution technology have been reaped and the possibility of further raising yield frontiers is uncertain. Low current world rice prices reduced incentives for public investments in irrigation and technological research throughout the 1980s. Yet, population growth rates in Asia are projected to increase at near 1.8% (United Nations). Even if the income elasticities for rice were low, it is not clear that the supply-demand outlook warrants the implications drawn by the authors.

[Received March 1990; final revision received July 1990.]

References

- Alderman, Harold. *The Effect of Food Price and Changes on the Acquisition of Food by Low-Income Households*. Washington DC: International Food Policy Research Institute, 1986.
- Asian Development Bank (ADB). *Evaluation of Rice Market Intervention Policies*. Manila, 1988.
- . *Key Indicators of Developing Member Countries of ADB*. Manila, various issues.
- Bouis, Howarth E. *Projects for Rice Supply and Demand Balances in Asia*. Final report submitted to Rockefeller Foundation. Washington DC: International Food Policy Research Institute, 1989.
- . "Rice in Asia: Is It Becoming a Commercial Good?" Working Paper. Washington DC: International Food Policy Research Institute, June 1990.
- Brown, William G., and Bruce R. Beattie. "Improving Estimates of Economic Parameters by Ridge Regression with Production Function Applications." *Amer. J. Agr. Econ.* 57(1975):21–32.
- David, Cristina C. "Global Rice Situation." Paper presented at the International Rice Research Conference, Beijing, China, Sep. 1987.
- Ito, Shoichi, E. Wesley F. Peterson, and Warren R. Grant. "Rice in Asia: Is It Becoming an Inferior Good?" *Amer. J. Agr. Econ.* 71(1989):32–42.
- International Monetary Fund (IMF). *International Financial Statistics, 1989*. Washington DC, 1989.
- Levine, Gil, R. Barker, M. Rosegrant, and M. Svendsen. "Irrigation in Asia and the Near East in the 1990's: Problems and Prospects." *Dep. Agr. Econ.*, Cornell University, 1988.
- Pardey, Philip G., and Johannes Roseboom. *ISNAR Agricultural Research Indicator Series*. Cambridge: Cambridge University Press, 1989.
- Quisumbing, Ma. Agnes R. "Estimating the Distributional Impact of Food Market Intervention Policies on Nutrition." Ph.D. thesis, University of the Philippines, 1985.
- U.S. Department of Agriculture (USDA), Foreign Agriculture Service. *Foreign Agriculture Circular, Grains*. Washington, DC, various issues.
- "World Population Prospects 1988." New York: United Nations, Dep. Int. Econ. and Soc. Affairs Population Stud. No. 106, 1989.

Rice in Asia: Is It Becoming a Commercial Good? Comment

Howarth E. Bouis

Ito, Peterson, and Grant (IPG) have recently shown that rice income elasticities are highly negative in many Asian countries and only slightly positive and declining over time in other Asian countries. Their results, if accurate, have obvious and important implications for how Asian governments, international lending agencies, and the CGIAR system should allocate resources in anticipation of weak demand and continued low international prices for rice.

This comment argues that IPG fail to consider the effect on demand for rice of decisions by semisubsistence farmers to produce rice, which leads to downwardly biased time-series estimates of aggregate national income elasticities. That is, demand effects that IPG attribute to trend increases in aggregate per capita income may largely reflect instead (a) a concurrent rural to urban migration pattern, (b) the decreasing importance over time of aggregate rice production within some Asian countries, and (c) a declining share of semisubsistence production within the rice sectors themselves.

Income Elasticities Derived from Expenditure Surveys

A useful starting point for arguing that production and consumption decisions for staple foods cannot be analyzed separately for Asian countries is to note that rice income elasticities derived from expenditure surveys are nearly always positive and in some instances considerably above zero. For example, Timmer and Alderman estimate an aggregate national income elasticity for rice for Indonesia of about 0.4 from household expenditure survey data collected in 1976. Pitt derives an estimate for Bangladesh of about 1.0 for 1973–74, and Trairatvorakul derives an estimate for Thailand of 0.13 for 1975–76 (see Alderman, table 12, for estimates for other countries).¹ The corresponding IPG estimates are 0.19, –0.04, and –0.25 for Indonesia, Bangladesh, and Thailand, respectively, for the survey years cited above.

The same comparison shows a similar result for a developed economy. IPG derive an estimate of –0.71 for the income elasticity of demand for rice in Japan for 1984, where household income averaged about

\$40,000 per year for a family of four.² Intuitively, such a high negative elasticity implies that budgets are so constrained that Japanese households consume rice in large amounts out of rather dire necessity. If incomes double and households reduced rice consumption by 70%,³ what (calorie dense) foods would these households eat instead which they could not already afford? In table 1 rice expenditure data presented by expenditure group for Japan suggest that higher income groups do not drastically reduce rice consumption as incomes increase.

How does one reconcile declining per capita rice consumption over time with such food expenditure information? Table 2 summarizes rice consumption data, disaggregated by expenditure quartile from food expenditure surveys in six less developed Asian countries. Controlling for urban and rural populations, per capita consumption of rice increases monotonically with income, except for Thailand where the arc income elasticity is close to zero. Comparing rice consumption levels within the same expenditure quartile, rural consumers consume more rice than urban consumers. The most pronounced and consistent difference between urban and rural consumption across income groups occurs for Thailand. IPG estimate that the aggregate income elasticity for rice for Thailand had fallen to –0.44 by 1985.

Explanations for Rural-Urban Demand Patterns

Lower rice consumption in urban versus rural areas has several possible explanations. First is the explanation put forth by IPG (although not in this particular context) that urban incomes are higher and that higher income consumers prefer to substitute more preferred foods for rice. However, this explanation does not go very far in explaining the large differences across urban and rural populations within the same expenditure quartiles indicated in table 2, nor is it supported by the trend in rice consumption across income quartiles, controlling for urban/rural.

Second, assume that two persons with the same set of demand preferences and levels of income and fac-

² GNP per capita in 1984 was US \$10,580 at current prices (World Bank, p. 340–41).

³ IPG cannot appeal to lower income elasticities closer to zero at, say, a household income of \$80,000 because their estimates appear always to become increasingly negative at higher income levels.

Howarth E. Bouis is a research fellow at the International Food Policy Research Institute, Washington, D. C.

¹ The estimate cited for Pitt is for rural households only.

Table 1. Per Capita Rice Expenditures, by Expenditure Group for Japan, 1983

Expenditure Group ^a	Nationwide		Worker's Households	
	Rice Expenditure ^b	Percent of Sample	Rice Expenditure ^b	Percent of Sample
Up to 999	2,032	0.01	1,912	0.00
1,000-1,499	1,764	0.02	1,630	0.01
1,500-1,999	1,772	0.04	1,524	0.02
2,000-2,499	1,647	0.07	1,471	0.05
2,500-2,999	1,593	0.08	1,456	0.07
3,000-3,499	1,482	0.10	1,393	0.10
3,500-3,999	1,532	0.10	1,409	0.11
4,000-4,499	1,522	0.10	1,467	0.11
4,500-4,999	1,587	0.09	1,480	0.10
5,000-5,499	1,578	0.07	1,498	0.08
5,500-5,999	1,660	0.06	1,678	0.07
6,000-6,499	1,786	0.05	1,732	0.06
6,500-6,999	1,747	0.04	1,716	0.05
7,000-7,499	1,598	0.03	1,597	0.05
7,500-7,999	1,643	0.02	1,649	0.03
8,000-8,999	1,692	0.04	1,663	0.04
9,000-9,999	1,688	0.03	1,645	0.03
10,000 and over	1,736	0.05	1,762	0.04

Source: Statistics Bureau, Prime Minister's Office, annual report on the family income and expenditure survey, as reported in Food and Agriculture Organization, review of food consumption surveys, 1985.

^a Thousand yen per household per year.

^b Yen per capita per month.

Table 2. Weekly Per Capita Consumption of Milled Rice, By Country, By Expenditure Quartile, By Urban and Rural Populations

Country	Urban or Rural	Expenditure Quartile ^a			
		1	2	3	4
(kilograms)					
Bangladesh	Urban	1.60	1.90	2.15	2.41
	Rural	1.45	1.98	2.51	3.13
India	Urban	1.08	1.34	1.36	1.36
	Rural	1.13	1.67	1.91	1.97
Indonesia	Urban	1.82	2.10	2.23	
	Rural	1.77	2.56	2.97	
Pakistan	Urban	0.15	0.20	0.24	0.29
	Rural	0.22	0.33	0.30	0.37
Sri Lanka	Urban	1.30	1.64	1.91	1.96
	Rural	1.63	2.09	2.52	2.88
Thailand	Urban	2.17	2.16	2.16	2.22
	Rural	2.92	3.13	3.00	2.89

Data sources (see also Bouis 1989b for other details): Bangladesh: *Household Expenditure Survey of Bangladesh 1973-74*, Bangladesh Bureau of Statistics; data are disaggregated by household expenditure quartile instead of per capita expenditure quartile. India: *Second Quinquennial Survey on Consumer Expenditure*, undertaken by the Department of Statistics from July 1977 to June 1978. Indonesia: Data are taken from SUSENAS 1981 expenditure surveys, as reported in Rosegrant et al. (1987). Pakistan: *Household Income and Expenditure Survey 1984-85*, conducted by the Federal Bureau of Statistics. Sri Lanka: Data are taken from the *1980/81 Labour Force and Socio-Economic Survey* conducted by the Bureau of Census and Statistics, as cited in Sahn (1984). Thailand: Data taken from the *1975/76 Socioeconomic Survey* conducted by the National Statistics Office, as cited in Trairatvorakul (1984); only calorie intake information is provided for various food groups; a calorie conversion rate of 3,500 calories per kilogram was used to compute rice consumption per kilogram.

^a Tercile in the case of Indonesia.

ing the same set of prices differ only in the activity levels necessary for earning that income. Energy requirements will be greater for the person engaged in the more active occupation. Under a reasonable set of assumptions (Bouis 1989a, 1990), the more active person will spend more for food than the less active person; more important, his/her diet will be more staple based because staples are relatively inexpensive sources of calories. If rural occupations require a greater energy expenditure than urban occupations, one would expect a priori to observe more staple consumption by rural populations.

Third, rice is cheaper to purchase in rural areas, in particular by rice producers themselves. Farmers who grow rice and consume out of own production avoid retailing and other marketing costs in staple purchases. Apart from greater risks associated with growing nonfood crops, such a cost saving may be an important additional incentive for semisubsistence farmers to grow staples. (See Bouis and Haddad 1990a, chaps. 4 and 7, for an analysis of this phenomenon for a specific group of farm households in the Philippines.)

In order to understand structural changes in demand behavior for rice over time in countries where a significant proportion of rice consumption comes from own production (all of the countries studied by IPG would qualify under these criteria at the start of the time-series data used, with the possible exception of Singapore), it is important to study three groups of consumers: (a) urban, (b) rural rice producers, and (c) rural nonrice producers. Thus, it is useful to write:

$$(1) \quad D_N = \theta_U D_U + \theta_{Rr} D_{Rr} + \theta_{Ro} D_{Ro},$$

where D is per capita consumption, θ is proportion of population, N is national total population, U is urban population, R is rural population, r is semisubsistence and commercial rice producers (these households may grow crops in addition to rice), o is producers of crops other than rice (these households grow no rice at all), and

$$(2) \quad \theta_u + \theta_{Rr} + \theta_{Ro} = 1.$$

Approaching analysis of national aggregate consumption in this way, it is conceivable that significant decreases in aggregate consumption could occur over time if consumption levels differed significantly such that, controlling for income,

$$(3) \quad D_U < D_{Ro} < D_{Rr}$$

and if there was significant rural to urban migration and/or rice production declined in importance in the rural sector. While increasing per capita incomes are usually associated with structural change in developing economies (i.e., rural to urban migration and a switch from subsistence to commercial agriculture), at some point these processes run their course (e.g., Japan and Taiwan). It then would be misleading to associate further increases in income with the same reduction in consumption of rice.

Correlations between Changes in Per Capita Rice Consumption and Structural Changes

Table 3 ranks the countries cited by IPG by percentage change in per capita rice consumption from 1961–65 to 1981–85.⁴ Also given in the table are percentage changes in rice production, agricultural gross domestic product, and total agricultural population, and change in percentage of the population which is urban.

The first relationship to note from table 3 is that rice production growth rates in Malaysia, Nepal, and Thailand (which IPG designate as Group I countries, experiencing substantial declines in per capita consumption of rice) were much lower than growth rates in agricultural gross domestic product in these countries. Rice production growth rates in these three countries were at or below the median as compared with rice production growth rates in other countries in South and Southeast Asia. Countries with the highest percentage increases in per capita rice consumption had the highest rice production growth rates, which tended to exceed growth rates for agricultural gross domestic product. Rice production became relatively more important within China and Indonesia, where the increases in per capita consumption of rice were highest.

Table 3 also indicates increased urbanization all over Asia. However, the trend in population densities in rural areas exhibited two distinct patterns as this urbanization occurred. Population densities in rural areas declined in Japan, South Korea, and Taiwan. Presumably as this occurred, rice farms increased in size, semisubsistence production gave way to commercial production, and marketed surplus as a percentage of total rice production greatly increased. These changes contrast with increased land pressures and production of rice on smaller operational units in several countries whose agricultural populations increased by 45% or more from the early 1960s to the mid-1980s. If yields had not increased so substantially in most of these countries, marketed surplus (as a percent of total production) likely would have fallen.

With the information in table 3 as background, consider the following implicit equation:

$$(4) \quad \% \Delta D_N = f(\% \Delta RCPRD - \% \Delta AGGDP, \% \Delta AGPOP, \Delta PCURB),$$

where Δ is change, $\% \Delta$ is percentage change, D_N is aggregate national per capita consumption of rice, $RCPRD$ is rice production, $AGGDP$ is agricultural gross domestic product, $AGPOP$ is agricultural population, and $PCURB$ is percent of total population which is urban. If $D_{Ro} < D_{Rr}$, then as rice becomes less important in total agricultural production, *ceteris paribus*, national aggregate consumption declines. Because data for D_{Ro} , D_{Rr} , θ_{Ro} , and θ_{Rr} are lacking, $\% \Delta RCPRD - \% \Delta AGGDP$ is a rough proxy (in con-

⁴ Singapore and Burma are excluded because of lack of data.

Table 3. Percentage Change in Rice Consumption, Rice Production, Agricultural Gross Domestic Product, Agricultural Population, and Change in Percent Urban, by Country, 1961 to 1985

Country	Percent Change in Per Capita Rice Consumption 1961-65 to 1981-85 ^a	Percent Change in Total Rice Production 1961 to 1985 ^b	Percent Change in Agricultural Gross Domestic Product 1967 to 1986 ^c	Percent Change in Agricultural Population 1961 to 1985 ^d	Change in Percent Urban ^e
Taiwan	-39	6	45	-24	29
Japan	-29	-8	-3	-66	22
Malaysia	-17	63	127	6	26
Nepal	-17	27	49	69	2
Thailand	-14	77	143	47	17
India	-3	63	54	53	6
Bangladesh	1	46	55	61	14
Sri Lanka	4	142	88	49	4
South Korea	5	54	85	-22	29
Philippines	12	102	116	54	11
China	33	160	130	35	12
Indonesia	47	166	133	17	22

^a Ito, Peterson, and Grant, table 1.^b IRRI, table 1; adjusted to reflect 20-year growth period.^c World Bank, 1988 world tables; adjusted to reflect 20-year growth period.^d IRRI, table 80.^e IRRI, table 80.

junction with $\% \Delta AGPOP$ discussed below) for the percentage change in the contribution of semisubsistence production of rice to total agricultural production. In the regression results presented in table 4, this variable is multiplied by the percent of the population which is rural at the beginning of the period, since the effect of $\% \Delta RCPRD - \% \Delta AGGDP$ on $\Delta \% D_N$ will be greater the larger is the rural sector.

$\% \Delta AGPOP$ is included as a control for farm size. *Ceteris paribus*, as farm size declines (as $\% \Delta AGPOP$ increases), there is a greater prevalence of low-income rural households producing rice for their subsistence needs, both to increase food security and to avoid paying marketing margins for their rice consumption. Finally, $\Delta PCURB$ controls for differences between urban and rural consumption of rice. If D_u

$< D_{ro}$, then the expected sign of the coefficient for $\Delta PCURB$ is negative. However, in most countries average urban incomes are higher than average rural incomes; if rice is not an inferior good, then controlling for production effects on consumption, the expected sign is positive.

The results of estimating (4) are shown in table 4. The rough-and-ready nature of the proxy variables used in the regression estimations, the small sample size, and the lack of variables which account for changes in prices, all provide admittedly compelling reasons for caution in interpreting the correlations shown in table 4. Still, the correlations are strong, have the expected signs, and explain much of the variance across countries. The positive coefficient on $\Delta PCURB$ suggests that higher income urban house-

Table 4. Results of Estimating Equation (4)

Sample/Variable	*Coefficient	t-*statistic
Full sample (12 countries)		
$(\% \Delta RCPRD - \% \Delta AGGDP) \%RURAL61^a$	0.84	4.0
$\% \Delta AGPOP$	0.60	3.0
$\Delta PCURB$	2.37	2.6
$R^2 = 0.71$		
Full sample, excluding Taiwan (11 countries)		
$(\% \Delta RCPRD - \% \Delta AGGDP) \%RURAL61^a$	0.83	5.7
$\% \Delta AGPOP$	0.57	4.2
$\Delta PCURB$	2.76	4.3
$R^2 = 0.84$		

Note: Dependent variable is percentage change in per capita consumption of rice, 1961-65 to 1981-85, as cited in table 3; data for all explanatory variables are cited in table 3 also, except for $\%RURAL61$.

^a $\%RURAL61$ is percent of total population which is rural in 1961 (IRRI, table 80).

holds do consume more rice than rural households, once production effects on consumption are controlled.

Conclusions

It is beyond the scope of this comment to provide conclusive evidence that the decline in rice consumption in some Asian countries (and increases in others) is largely the result of structural changes in income generation and not to the amount of income that is earned, although the two can be correlated over time. The main benefit from estimating (4) is to show that the aggregate country data at least do not contradict, indeed are very consistent with, two propositions that merit further research—first, rice consumption in Asia does not decline rapidly with income per se, and second, national per capita rice consumption levels have been and in many countries still are closely related to rice production relationships that determine marketed surplus (farm size, irrigation, and road infrastructure, adoption of modern technologies).

In countries where semisubsistence rice production is still important, aggregate national rice consumption may continue to decline as agriculture becomes more commercial and especially as rural population densities decline. However, predicting the existence and rate of such a decline will depend on a much better understanding of the effects of the commercialization process and of reduced work activity levels (which presumably accompany economic development) on demand for specific foods, and for staples in particular.

More specifically, after controlling for income, how much of the difference between urban and rural consumption (shown in table 2) can be explained by (a) generally lower food prices in rural areas, (b) lower implicit rice prices which rice producers pay for consumption of own-production, and (c) by higher work activity levels in rural areas? Are there factors related to expenditure survey techniques that lead to gross overestimates of rice consumption for rural consumers, as compared with urban consumers (Bouis and Haddad 1990b)?

Can supporting evidence be found that the substantial declines in per capita consumption of rice over time in Japan and Taiwan resulted from a decline in semisubsistence rice production? Why has rice consumption not fallen in South Korea, as apparently happened in Japan, Singapore, and Taiwan? Did these substantial declines in consumption in fact not occur in some instances? Are some of the time-series data used by IPG derived from unreliable production and stock information?⁵

⁵ The validity of the 222 kilograms per capita consumption of milled rice cited for Burma for 1981–85 can be questioned on a priori grounds. Such a level of consumption corresponds to a calorie availability of about 2,800 calories per adult equivalent per day, which is about 10% above the recommended intake for the Philippines from all foods. Similarly, the figure of 191 kilograms per capita for Thailand for 1961–65 is quite high.

Answers to these questions are crucial for planning investments for agricultural research and irrigation. Because of the substantial time lags in realizing investment benefits, if these investments are reduced or delayed because of overly pessimistic demand projections, rice prices may rise unnecessarily with adverse consequences for the poor in Asia, urban and rural, who are net consumers of rice.

[Received June 1990; final revision received August 1990.]

References

- Alderman, H. *The Effect of Food Price and Income Changes on the Acquisition of Food by Low-Income Households*. Washington DC: International Food Policy Research Institute, 1986.
- Bangladesh Bureau of Statistics, Statistics Division, Ministry of Planning. *A Report on the Household Expenditure Survey of Bangladesh 1973–74, vol. 1*. Dacca, 1978.
- Bouis, H. E. "A Food Demand System Based on Demand for Characteristics: If There Is 'Curvature' in the Slutsky Matrix, What Do the Curves Look Like and Why?" Paper presented at the 33rd Conference of the Australian Economics Society, Christchurch, New Zealand, 1989a.
- . "Evaluating Demand for Calories for Urban and Rural Populations in the Philippines: Consequences for Nutrition Policy Under Economic Recovery." *World Develop.* 18(1990):281–99.
- . "Prospects for Rice/Supply Demand Balances in Asia." Washington DC: International Food Policy Research Institute, 1989b.
- Bouis, H. E., and L. J. Haddad. *Agricultural Commercialization, Nutrition, and the Rural Poor: A Study of Philippine Farm Households*. Boulder CO: Lynne Rienner Publishers, 1990a.
- . *Are Estimates of Calorie-Income Elasticities Too High: A Reevaluation of the Plausible Range*. Washington DC: International Food Policy Research Institute, 1990b.
- Federal Bureau of Statistics, Statistics Division, Government of Pakistan. *Household Income and Expenditure Survey 1984–85*. Pakistan, undated.
- Food and Agriculture Organization of the United Nations. *Review of Food Consumption Surveys—1985*. FAO Food and Nutrition Pap. No. 35. Rome, 1986.
- International Rice Research Institute. *World Rice Statistics 1987*. Los Banos, Philippines, 1988.
- Ito, Shoichi, E. Wesley F. Peterson, and Warren R. Grant. "Rice in Asia: Is It Becoming an Inferior Good?" *Amer. J. Agr. Econ.* 71(1989):33–42.
- National Sample Survey Organization, Department of Statistics, Government of India. *Report on the Second Quinquennial Survey on Consumer Expenditure, thirty-second round July 1977–June 1978*. New Delhi, 1984.
- Pitt, M. M. "Food Preferences and Nutrition in Rural Bangladesh." *Rev. Econ. and Statist.* 65(1983):105–14.
- Rosegrant, M., F. Kasryno, L. Gonzales, C. Rasahan, and Y. Saefudin. *Price and Investment Policies in the Indonesian Food Crop Sector*. Washington DC: International Food Policy Research Institute, and Bogor,

- Indonesia: Center for Agro-Economic Research, 1987.
- Sahn, D. *Food Consumption Patterns and Parameters in Sri Lanka: The Causes and Control of Malnutrition*. Washington DC: International Food Policy Research Institute, 1984.
- . "The Effect of Price and Income Changes on Food-Energy Intake in Sri Lanka." *Econ. Develop. and Cultur. Change* 36(1988):315–40.
- Timmer, C. P., and H. Alderman. "Estimating Consumption Parameters for Food Policy Analysis." *Amer. J. Agr. Econ.* 61(1979):982–87.
- Trairatvorakul, P. *The Effects on Income Distribution and Nutrition of Alternative Rice Price Policies in Thailand*. Washington DC: International Food Policy Research Institute Res. Rep. No. 46, 1984.
- World Bank. *World Tables, 1988–89 Edition*. Washington DC, 1989.

Rice in Asia: Is It Becoming an Inferior Good? Reply

Shoichi Ito, E. Wesley F. Peterson, and Warren R. Grant

Given the prominence of rice in Asian economies as well as the importance of Asia for world rice markets, it is not surprising that the results we reported in our article (Ito, Peterson, and Grant) were somewhat controversial. We concluded that declining income elasticities of demand for rice in Asia mean that predictions based on conventional elasticity estimates are likely to overestimate total demand. Although the authors of the two comments reach their conclusions using very different approaches, they appear to agree that our predictions underestimate future rice demand in Asia. This debate is important because certain policy decisions in Asia, as well as in other nations exporting or importing rice, may depend on how actual Asian rice demand evolves. In our reply, we respond to each of the comments in turn.

Reply to Huang, David, and Duff

Huang, David, and Duff raise a number of issues concerning data and method. In particular, they argue for the following procedures: (a) use of private consumption expenditures expressed in a common currency rather than per capita gross domestic product (GDP) expressed in national currencies, (b) use of domestic price data instead of world price data, and (c) the use of prior information in specifying the ridge regression. Using these procedures, they reestimate our model and obtain results that differ from the results we reported. However, their results do not differ greatly from those we reported. While the income elasticities estimated by Huang, David, and Duff decline more slowly than those we estimated and remain positive in two countries where our estimates indicate that rice has become an inferior good, the general tendencies predicted in our study are replicated in theirs.

There will always be differences in elasticity estimates derived from different research approaches and data sets (Gardiner and Dixit). Choosing between two sets of estimates is complicated in this case because, due to the use of ridge regression, the diagnostic statistics associated with each model are not reliable. Although the estimated standard errors in our

model are substantially smaller relative to their corresponding coefficients than is the case for the model reported by Huang, David, and Duff, this cannot be taken as an indicator of statistical superiority because of the biases introduced by using ridge regression. Price has developed a formal procedure for comparing two models of this nature, but such a test requires more data, space, and time than we have available. However, the challenge to our results presented by Huang, David, and Duff is based critically on the procedural points they raise. In the absence of statistical indicators, an examination of the validity of these points is a reasonable way to approach the evaluation of their criticisms.

The first procedural point raised by Huang, David and Duff concerns the data series used to represent income. Huang, David, and Duff argue that per capita private consumption expenditure (PCE) is a better measure of income than gross domestic product (GDP) because it does not include government expenditure which, they suggest, is likely to grow more rapidly than the private consumption components of GDP. Although it may indeed be the case that, over time, a greater proportion of GDP is accounted for by the governments of these countries, this part of the argument is an empirical question for which the authors provide no support. If the change in composition of GDP is not great, it is likely that the two series will be highly correlated, so it makes little difference which is used. Correlation coefficients between PCE and GDP were computed for all the countries except Bangladesh, Nepal, the PRC, and Taiwan for which data were not available from the International Monetary Fund (IMF). These coefficients as well as the results of regressing PCE on GDP are shown in table 1. The results indicate that the two series are very highly correlated (most of the correlation coefficients are either 0.999 or 0.998), and we conclude that it probably makes little difference which series is chosen.

In addition to criticizing the data series chosen to represent the income variable, Huang, David, and Duff suggest that income should be expressed in a common currency rather than in individual national currencies. According to the authors, their income variable, *PCE*, is expressed "in real terms (deflated by consumer price index) . . . in US dollars converted by the purchasing power parity (PPP) exchange rate for 1980." The PPP exchange rate should account for relative inflation rates between the individual countries and the *numéraire* country, presumably the United

Shoichi Ito is an assistant professor, Department of Information Science, Tottori University, Japan; Wes Peterson is an associate professor, Department of Agricultural Economics, University of Nebraska; and Warren Grant is an adjunct professor, Department of Agricultural Economics, Texas A&M University.

Table 1. Statistical Relationship between Gross Domestic Product (GDP) and Private Consumption Expenditure (PCE) in Asian Countries, 1961–85

Country	Regression Coefficients ^a	R ²	Correlation Coefficient ^b
Japan	0.592 (0.005)	0.998	0.999
Malaysia	0.493 (0.005)	0.997	0.999
Singapore	0.425 (0.010)	0.987	0.994
Thailand	0.652 (0.003)	0.999	0.999
India	0.671 (0.004)	0.999	0.999
South Korea	0.603 (0.007)	0.997	0.999
Sri Lanka	0.769 (0.008)	0.997	0.999
Burma	0.860 (0.010)	0.996	0.998
Indonesia	0.573 (0.006)	0.998	0.999
Philippines	0.736 (0.011)	0.995	0.998

Data Source: International Monetary Fund.

^a The regression coefficients are the slopes from equations with PCE as the dependent variable. The figures in parentheses are standard errors.

^b Correlation coefficient between PCE and GDP in each country.

States, so it is difficult to see why the income series should also be deflated by some unidentified consumer price index. In addition, the use of the PPP exchange rates introduces variables (e.g., the U.S. inflation rate) that may have little to do with domestic demand in these countries. The income variables used by Huang, David, and Duff may be appropriate for countries such as Singapore, where all of the rice consumed is imported. For the other countries where imports generally make up less than 10% of domestic consumption (except in Malaysia where they account for 20%), the distortions introduced by this conversion may be worse than the statistical problem signalled in their comment.

The second procedural point raised by Huang, David, and Duff concerns the price data. It is, of course, preferable to use domestic price data rather than world prices if domestic prices are accurately reported. The reason we chose to use world prices is that we were unable to locate reliable national price series for rice and wheat. Huang, David, and Duff apparently have access to such series, although they do not document their sources. The authors indicate that their domestic price series and the ratios they discuss in the text of their comment are taken from "individual country sources" (p. 4), none of which is cited in the references. Although government policies can drive a wedge between domestic and international prices, it is difficult to isolate the domestic

market completely from world market conditions. Even in Japan, with its highly protectionist rice policies, internal rice prices are not completely insulated from changes in world prices. The correlation coefficient between the Japanese government procurement price (deflated by the Japanese consumer price index) and the Thai rice export price converted to yen (also deflated by the Japanese consumer price index) and lagged one year is 0.76 over the period 1961–89 (based on data from Food Agency and IMF). This suggests that the use of world prices may not seriously undermine our results. Further, if the domestic prices obtained from individual country sources are not accurately reported, the statistical results could be worse than those based on the more reliable world price series taken as proxies for domestic prices.

The third procedural point raised by Huang, David, and Duff concerns the use of prior information in ridge regression. They suggest that the following criteria should be used in choosing the k -values for the ridge regression: (a) the k -values should be small; (b) they are expected to give rise to a negative coefficient for the inverse of the income variable; (c) the coefficients from the model should lead to elasticities that are close to the constant elasticities obtained from a double-log specification. The first criterion is not controversial. In specifying our model, we specified small k -values (0.00002, 0.05, and 0.1 compared with 0.002, 0.01, and 0.0005 in Huang, David, and Duff). With respect to the second criterion, we also expected the coefficients for the inverse-income variables to be negative in line with our expectation that income elasticities in Asia are declining as incomes increase, as we stated in the text. We did not constrain the coefficient estimates to be negative and, apparently, Huang, David, and Duff did not do so either because they report positive coefficients for the inverse-income variable in two countries, Bangladesh and South Korea. In terms of including the "prior information" embodied in the first two criteria suggested by Huang, David, and Duff, there does not appear to be much difference between the two models.

The third criterion really is not very useful. It does not provide a guide for determining what constitutes an estimate that is "not very far from the constant elasticities . . ." (p. 4). The justification for attaching particular weight to constant elasticity estimates appears to be statistical (see their footnote 5). There is no theoretical reason for taking elasticity estimates from a double-log specification as benchmarks for the empirical analysis. The double-log specification does not satisfy the adding-up condition of theoretical demand analysis and, thus, makes no more sense than some other specification as a reliable indicator of the order of magnitude of the elasticities (Phlips). The justification in footnote five would apply equally well to an ordinary linear model.

In addition, applying this criterion to the results presented in the authors' table 5 does not unambiguously favor the k -values they selected. For example, consider the estimates for 1975. Only four of the es-

estimated elasticities in the column under k^* are closer to the constant elasticity estimates than estimates recorded in other columns. Seven of the estimates based on a k -value of 0.00001 are closest to the constant elasticities, with one each in the other columns. In fact, out of thirty-nine possibilities in table 5, only twelve of the estimates in the k^* column are "better" in the sense of being closer to the constant elasticities. One wonders exactly how the "prior information" contained in the constant elasticities was used to select the preferred k -values.

The comment by Huang, David, and Duff is based primarily on these procedural points and a few unsupported assertions. We have responded to the procedural points and leave it to the reader to decide whether the criticisms of Huang, David, and Duff are telling. On the whole, their analysis appears to be a somewhat "ad hoc" attempt to generate outcomes consistent with certain preconceptions. For example, the authors argue that negative income elasticities are implausible in low income countries. This is an empirical issue, and it is misleading to assume at the outset that a commodity can never be an inferior good in low-income countries. Daly et al. reported a negative income elasticity for rice in rural Vietnam in 1973, a time when that region would certainly have qualified as "low income." Another example of an unsupported assertion is the suggestion by Huang, David, and Duff that income elasticities in Thailand can be "expected to become negative only sometime in the 1980s." According to their own estimates presented in their table 3, the Thai income elasticity became negative in 1967.

Finally, Huang, David, and Duff argue that, based on their results, the countries with negative income elasticities account for only 10% of Asian rice consumption. They speculate that income elasticities for rice in most Asian countries will remain positive throughout the 1990s and that rice demand in the region will remain strong. In another study (Peterson, Lan, and Ito), we obtained estimates indicating that the income elasticity of demand for rice in the People's Republic of China became negative in the early 1980s. The income and price data used in this study were from Chinese sources, and the analysis was based on time-series estimates of the log-inverse-log model. These results are not subject to any of the criticisms Huang, David, and Duff raise because the data meet all of their recommendations, ridge regression was not used, and the Chinese account for more than one-third of total world rice consumption.

Reply to Bouis

In the other comment, Bouis argues that changing patterns of rice consumption in Asia may be due primarily to structural changes in Asian economies, notably those related to rural migration and the commercialization of rice production. These results actually complement ours in that they provide insights into

the processes behind declining per capita rice consumption in some countries. Bouis's results appear to support his conclusion that aggregate declines in rice consumption may occur as "agriculture becomes more commercial and especially as rural population densities decline." This conclusion can be linked to our results if it is true that income increases in Asia are associated with agricultural commercialization and declines in rural population densities. While this association seems plausible, an empirical test is still needed. At any rate, Bouis's focus on country-specific explanations for aggregate changes is a useful contribution to the discussion of Asian rice demand.

Bouis does recognize some limitations in his analysis. We would add a few procedural and conceptual points to his list. In discussing rice consumption, Bouis relies primarily on expenditure data. Rice expenditure is not the same as rice consumption because different qualities of rice mean that there will be different prices at any given time. In Japan, for example, rice prices range from around 300 yen per kilogram to over 600 yen per kilogram depending on the quality (Ito, Wailes, and Cramer). Ito, Wailes, and Cramer have shown that higher-income consumers tend to purchase higher-quality rice rather than larger quantities. The figures in Bouis's table 1 do not reflect shifts of this nature, and his inferences concerning the quantities consumed by higher-income consumers may be misleading. Similar comments may apply to his second table depending on whether the quantities shown were obtained by dividing expenditures by average prices or whether data on quantities consumed were also collected in the expenditure surveys cited.

Bouis's empirical analysis is based on an equation relating changes in aggregate rice demand to a set of proxies that are supposed to measure changes in semisubsistence rice production and farm size, controlling for differences in rural and urban consumption. This equation does not appear to have a basis in economic theory because prices and incomes, variables normally associated with the theory of demand, are not included. Real retail rice prices are shown for most of the countries under discussion in table 2. In many cases these prices fell considerably between 1961 and 1981, a fact which may be important in accounting for consumption patterns. Even if one accepts Bouis's specification, it is not clear that the proxies used in the statistical exercise are good measures of the target variables. For example, the first variable is supposed to measure the "contribution of semisubsistence production of rice to total agricultural production." To measure this, however, Bouis uses total rice production and agricultural GDP normalized on the percentage of the population that was rural in 1961. This might make sense as a proxy if all changes in rice production took place on semisubsistence farms. This may be true in some countries but probably not in all of them.

In motivating his comment, Bouis suggests that we would explain changing consumption patterns by claiming that "urban incomes are higher and that higher

Table 2. Comparison of Real Retail Prices of Milled Rice in Asian Countries between 1961 and 1981

Country	1961 ^a	1981 ^b	Real Price of 1961 Rice ^c	Change ^d
Bangladesh (taka/ton)	796 (11.5) ^f	7,613 (116.2)	8,043	(%) -5.4
Burma (kyat/ton)	371 (27.3)	1,286 (100.3)	1,361	-5.5
India (rupee/ton)	608 (26.1)	2,272 (113.0)	2,632	-13.7
Indonesia (rupiah/kg)	44 (20.0)	212 (112.2)	247	-14.2
Japan ^e (yen/60 kg)	4,351 (24.2)	16,381 (104.9)	18,838	-13.0
South Korea (won/kg)	23 (6.6)	811 (121.3)	423	91.7
Malaysia (M\$/ton)	463 (51.4)	990 (109.7)	989	0.1
Philippines (peso/ton)	470 (15.4)	2,660 (113.1)	3,451	-22.9
Sri Lanka (rupee/t.)	1,277 (32.9)	6,250 (117.9)	4,576	36.6
Taiwan (NT\$/ton)	5,390 (29.3)	19,600 (116.3)	23,378	-16.2
Thailand (baht/ton)	2,320 (33.9)	7,900 (118.6)	8,125	-2.8

Sources: IIRI, IMF, Republic of China, and Food Agency (MAFF, Japan).

^a The Indonesian and Japanese figures are for 1970 and 1960, respectively.

^b Figures for Taiwan are for 1979.

^c Calculated using the consumer price index figures and expressed in real 1981 values.

^d 1981 rice prices relative to real 1961 prices.

^e Brown rice.

^f Numbers in parentheses are consumer price indices, 1980 = 100.

income consumers prefer to substitute more preferred foods for rice." In fact, we make no claims with respect to urban incomes. Our study relates national consumption to national income, focusing on aggregate income elasticities of demand. Thus, Pitt's elasticity estimates for rural areas are not strictly comparable to ours. Rural and urban demand elasticities often differ, as we noted (see also Moon, KREI, and Daly et al.). Further, we would emphasize that our estimates, as in the case of any elasticity estimates, pertain to marginal changes, not the large change Bouis suggests in his thought experiment on Japanese consumption at the beginning of his comment.

reason to alter the overall conclusions we presented in our article. The results of our study on China (Peterson, Lan, and Ito) support and strengthen those conclusions, although we are more willing to defend the trends we identified in the two studies than the exact magnitude of the estimated income elasticities in specific countries at specific times. Bouis emphasizes the potential social costs of insufficient investment if future rice demand is underestimated. It is important to note that there would also be social costs if rice demand is overestimated and too much is invested in this sector at the expense of other economic activities.

[Received October 1990; no revision.]

Conclusion

The authors of both comments have raised interesting points. In the conclusion of our original article we commented on some of the limitations of the analysis. Huang, David, and Duff, and Bouis suggest other limitations and offer further recommendations for an improved understanding of Asian rice demand. Their contributions are useful in refining the issues related to this important topic and offering alternative visions of the changing markets for rice in Asia. After reflecting on their comments, however, we see no

References

- Daly, Rex F., R. G. Hoffman, F. Nelson, and H. Weingarten. *Agriculture in the Vietnam Economy: A System for Economic Analysis*. Washington DC: U.S. Department of Agriculture, Econ. Res. Serv., June 1973.
- Food Agency. *Beika ni Kansuru shiryō (Data Related to Rice Prices)*. Tokyo: Ministry of Agriculture, Forestry and Fisheries (MAFF), various issues.
- Gardiner, W. H., and P. Dixit. *The Price Elasticity of Ex-*

- port Demand: Concepts and Estimates*. Washington DC: U.S. Department of Agriculture, Econ. Res. Serv. Rep. No. AGES 860408, 1986.
- International Monetary Fund (IMF). *International Financial Statistics*, yearbooks and monthly issues. Washington DC, various dates.
- International Rice Research Institute (IRRI). *World Rice Statistics, 1985*. Manila, Philippines, 1985.
- Ito, Shoichi, E. Wesley F. Peterson, and Warren R. Grant. "Rice in Asia: Is It Becoming an Inferior Good?" *Amer. J. Agr. Econ.* 71(1989):32-42.
- Ito, Shoichi, Eric J. Wailes, and Gail L. Cramer. *Japan's Rice Markets and Policies*. Dep. Agr. Econ. and Rural Soc. Staff Pap. No. SP0190, University of Arkansas, 1990.
- KREI (Korea Rural Economics Institute). *Seoul-shi Migok Utong Gaesun Bang Hyang (Studies for Improving Rice Marketing in Seoul Area-Focusing on the Wholesale Market)*. Seoul, South Korea, 1984.
- Moon, Pal Yong. "The Evolution of Rice Policy in Korea." *Food Res. Inst. Stud.* 14(1975):381-402.
- Peterson, E. Wesley F., Lan Jin, and Shoichi Ito. "An Econometric Analysis of Rice Consumption in the People's Republic of China." *Agr. Econ. J. Int. Assoc. Agr. Econ.*, in press.
- Philips, Louis. *Applied Consumption Analysis*. New York: North-Holland Publishing Co., 1983.
- Republic of China. *Taiwan Statistical Data Book 1985*. Taipei: Council for Economic Planning and Development, 1985.

The Efficiency of Alternative Policies for the EC's Common Agricultural Policy: Comment

Alison Burrell

In a recent article in this *Journal*, De Gorter and Meilke (DGM) compare the redistributive efficiency of agricultural price support policies under alternative assumptions about accompanying supply control measures. Using a long-run supply curve which is assumed invariant with respect to policy choice, DGM assess the consumer/taxpayer cost per ECU of total producer surplus under different policies and conclude that individual production quotas followed closely by a two-tier price policy are preferable to price reductions or a coresponsibility levy.

The main substance of this note is that DGM's failure to recognize the underlying structural problems of EC agriculture and to identify the true political target of income support measures renders their approach incapable of yielding a determinate ranking of these measures in terms of their medium-term efficiency and invalidates their policy recommendations.

EC agriculture still suffers from the problem of too many producers of inadequate economic size. In the mid-1980s, agricultural employment in EUR 12 was still over 8% of total civilian employment. About one-third of total agricultural labor input was used on the 38% of farms in the very small and small size groups (average size about 6 hectares), which produced only 8% of total output (Commission). This pattern holds for most individual outputs. In particular, 51% of all holdings growing cereals in 1985 in EUR 10¹ had less than 5 cereal hectares, and 80% less than 10 hectares (Eurostat).

Paradoxically, the structural weakness of European agriculture is both the product of past policy and a continuing political constraint on attempts to reduce CAP price support. Despite the increased emphasis on structural policies for withdrawing surplus resources from agriculture (early retirement, dairy cessation, land set-aside, land conversion to nonagricultural use), there is still considerable scope for the capture of economies of size.

Clearly, medium-term structural evolution in EC agriculture is endogenous with respect to the choice of policy. It follows that an appropriate analytical framework for comparing medium-term policy outcomes must explicitly recognize differential policy-induced shifts in the supply curve. Moreover, DGM's device of standardizing the outcomes of alternative

policies by holding constant the aggregate producer surplus for the sector may be theoretically acceptable, but it is unsatisfactory as a basis for policy recommendations in the EC context, where the political target of income support is the return to farm-owned assets of the marginal full-time farmer or per agricultural work unit. DGM's approach ignores the fact that policy-induced changes in the number of producers and the distribution of total supply over producers will differ between policies, resulting in different values of the political target for a given value of aggregate producer surplus.

Initially, a quota scheme freezes structural change. Its medium-term impact depends on the arrangements for quota transfer. The greater the constraints placed on the free transfer of quota between producers, the more slowly cost-reducing structural change can proceed. Attempts to counter this effect by increasingly expensive producer buy-out programs followed by quota redistribution to existing producers result in additional program costs which should be acknowledged in the comparative costing of policies.

On the other hand, if quota is freely marketable, structural change is more rapid than in the absence of quota markets, permitting greater improvements in the efficiency of resource use but at the expense of policy efficiency. Over time, an increasing proportion of income support is capitalized into the value of quota.² As quota sellers are bought out, they take capital out of the sector. Established producers or new entrants who buy quota do so from owned or borrowed capital. If the capital is borrowed, the claim on assets in the sector becomes more diffuse and some income support intended for producers is creamed off by lending institutions. If the capital is supplied by the farmer, this extra capital still has an opportunity cost and must earn a return comparable with alternative uses. In either case, the residual return to the traditional farmer-owned assets of land, labor, management, and productive capital (excluding quota) is squeezed. This can be visualized as an upward displacement of the medium-term supply curve in DGM's figure 1.

By contrast, both the two-tier price and the price reduction policies lower the price on marginal units and on some or all intramarginal units and so provide

Alison Burrell is a senior lecturer in agricultural economics, Wye College, University of London.

¹ EC less the two most recent entrants, Spain and Portugal.

² For instance, there has been significant, capitalization of income support into milk quota values in many EC member states since the scheme began in 1984, with quota prices reflecting the high marginal net return to expansion due to economies of size.

a greater stimulus to structural rationalization. Falling costs cause a more rapid downward displacement of the aggregate supply curve, so that over time the relative redistributive efficiency of these policies (as defined by DGM) changes. At the same time, producer numbers decline. It is the net effect of these changes which determines producer surplus per agricultural labor unit, according to which these alternatives should be ranked. It is impossible to determine the ranking from an analysis in which structural change is ignored. However, it seems unlikely on a priori grounds that the quota alternative would be optimal.

This view assumes that for the EC cereal sector and EC agriculture generally (a) there are significant potential size economies and (b) it is economically justified and politically realistic to ignore the costs to individuals of structural change. Regarding (a), we cite the poor size distribution of production referred to above, the existence of size-dependent technologies involving greater mechanization and specialization of management, and the copious literature (discussed recently by Castle, for instance) on the role of increasing size in the adoption of new technology.

It can also be argued that the economic and political climate is increasingly favorable to proposition (b). The age distribution in EC farming is much more heavily weighted toward the over-50 age groups than in the industrial or service sectors (Commission), and labor recruitment problems resulting from demographic changes are forecast in these sectors for the 1990s. More information is now available about the declining role of farming in the economics of many farm families. For instance, although three-quarters of holdings in the Federal Republic of Germany grow cereals, about half those in the field crops/mixed crops/crops and livestock categories have other gainful activity apart from farming (Eurostat). All this evidence suggests that the perceived social and political costs of structural rationalization may be lower now than at any time in the history of the CAP.

DGM's comparison of the terms-of-trade effects of these policies is also questionable. With exogenous world excess demand for EC exports, the terms-of-trade benefits are greatest for the quota program because it involves the largest reduction in the EC exportable surplus. However, these benefits may be short-lived. For example, in markets where world prices have been undermined by the export of EC surpluses (most notably dairy products and sugar) some of the world's least-cost exporting countries have curbed their export capacity in order to avoid further deterioration of the terms of trade against them. Short-run terms-of-trade benefits enjoyed by the EC following the introduction of a quota scheme could fall away as the volume of world trade is partly restored by these countries. Once again, it is a question of the medium-term endogeneity of supply conditions with respect to EC policy choice, but this time in other trading countries. Are these medium-term adjust-

ments incorporated (via lagged prices or otherwise) into DGM's function for world excess demand for EC exports?

Finally, two practical points should be stressed. First, DGM note that the dual role of wheat as both intermediate input (feed) and final product (food) would create problems in implementing a two-price scheme and hence their final recommendation of a quota scheme for this sector. However, under a quota system there would be the same tendency for feed wheat user prices to equal marginal production cost because cereal farmers with livestock production would use their own wheat (at cost) rather than buying it at the higher market price.³ Second, the administrative difficulty of imposing the enforcing individual quotas in a sector where half of all holdings (about 2.5 million farms in EUR 10) grow cereals (Eurostat) is probably insuperable.⁴

In conclusion, the long-term problems of EC agriculture must ultimately be solved by structural rationalization. DGM ignore the medium-term endogeneity of structural change and fail to incorporate into their model the structural impact of each policy. The current willingness of EC policy makers to increase their reliance on structural measures and on the targeting of assistance to particular groups of producers while at the same time taking a stronger line on support prices marks a step toward the decoupling of income support from market management and the renewal of the link between marginal supply and market conditions. A quota program is a backward move on this policy continuum. DGM's conclusions, skillfully argued though they are, are based on a false identity between the long-run properties of their model and the long-run realities of CAP reform.

[Received February 1990; final revision received August 1990.]

References

- Castle, Emery N. "Is Farming a Constant Cost Industry?" *Amer. J. Agr. Econ.* 71(1989):574-82.
- Commission of the European Communities. *The Agricultural Situation in the Community*. Brussels and Luxembourg, various years.
- De Gorter, Harry, and Karl D. Meilke. "Efficiency of Alternative Policies for the EC's Common Agricultural Policy." *Amer. J. Agr. Econ.* 71(1989):592-603.
- Eurostat. *Farm Structure, 1985 Survey: Main Results*. Brussels and Luxembourg, 1987.

³ I am indebted to Stefan Tangermann for this point.

⁴ Compare milk: only about 27% of EC farms have dairy herds and, given the nature of the product, the administrative framework for monitoring individual supplies already existed prior to quotas.

The Efficiency of Alternative Policies for the EC's Common Agricultural Policy: Reply

Harry de Gorter and Karl D. Meilke

Burrell argues that our welfare evaluations of alternative policy options considered by the European Community (EC) for the Common Agricultural Policy (CAP) fail to recognize policy-induced shifts in the supply curve. Burrell concludes that our approach is therefore "incapable of yielding a determinate ranking" and hence "invalidates" our policy recommendations. Although we analyzed four policy mechanisms and showed both theoretically and empirically that a production quota, under some circumstances, is inferior to a two-price plan, Burrell determines unequivocally that a quota is an inferior instrument.

Her conclusion is based on several arguments. First, nontransferable quotas induce economic inefficiency. This is true; but at no time in our original paper did we recommend that quotas, or any of the other three instruments for that matter, be implemented in an inefficient manner. We assumed transferable quotas.

A comprehensive evaluation of alternative implementation methods for each of the four policy instruments was beyond the scope of our paper. Nevertheless, we question Burrell's assertion that quotas will necessarily be implemented in a more inefficient manner than, for example, a two-price plan. To motivate the problem, consider Canada's production quota scheme for milk and compare it to the two-price plan used in the United States. Nontransferability of quotas between provinces in Canada induces deadweight costs additional to those calculated in our paper. Similar costs are also incurred currently in the EC with nontransferable milk quotas between countries.¹ However, Class I minimum prices in the United States are set in relation to Eau Claire, Wisconsin, so that the most deficit region in the United States is self-sufficient in milk production. This induces additional social resource costs resulting from distorted regional production patterns akin to those induced by nontransferable quotas in Canada and the EC (Public Hearing on Federal Milk Marketing Orders).² Indeed, the U.S. two-price plan has another ineffi-

ciency associated with it: farmers' production decisions are determined by the blend price rather than the lower Class II price. These inefficiencies of the U.S. program may well induce more additional social costs than a nontransferable quota scheme.³

Interestingly, a production quota scheme superimposed on a two-price plan can improve efficiency. The United Dairymen of Arizona, representing all dairy farmers in Arizona, unilaterally instituted production quotas for Class I milk in the 1970s (de Gorter and Schmitz). Any milk produced in excess of this quota received the Class II price. This effectively eliminated the blend pricing scheme facing Arizona's farmers and hence reduced output. Because the federal government of the United States fixes the Class I and II prices to consumers, the latter's welfare was unaffected. The welfare of Arizona farmers increased because the high Class I price was not diluted by the lower Class II price, and costs of overproducing at the blend price were not incurred. U.S. taxpayers benefitted because production was lower, and given fixed prices to consumers, resulting in lower costs in removing Class II milk products from the market. Production quotas were therefore a Pareto-superior move!⁴

The open-ended price support system of the EC with unlimited export restitutions also has inefficiencies that are ignored by Burrell and in our original paper. Since the late 1960s, producer prices in the EC have been anything but "common" as a result of green rates and member country subsidies which have resulted in the renationalization of agricultural subsidization (Runge and von Witzke). The incremental costs to a country of these subsidies are often less than the incremental benefits net of the increase in its share of the budget burden induced by this production increase. We call this phenomenon easy riding on the CAP, rather than free riding because it costs a little to gain. As a result, the social costs of the CAP are higher; but these costs could be avoided with a production quota scheme that would limit the production effects of member country subsidies.

The above discussion illustrates the impossibility

Harry de Gorter is an assistant professor, Department of Agricultural Economics, Cornell University; Karl Meilke is a professor, Department of Agricultural Economics and Agribusiness, University of Guelph.

¹ Milk quotas are transferable between farms in most provinces in Canada just like they are in most EC countries.

² Rucker, Thurman, and Sumner recently analyzed the welfare costs of nontransferability of tobacco and peanut quotas across county lines and found the costs to be minimal.

³ Another inefficiency induced by the blend pricing scheme in the United States which would not have occurred with quotas was the unnecessary conversion of grade B milk production enterprises to grade A enterprise.

⁴ The irony is that the quotas were disbanded after being deemed unconstitutional in a lawsuit brought forward by the Department of Justice. See de Gorter and Schmitz for an interesting discussion.

of arguing categorically that quotas are an inferior instrument because of nontransferability.

Second, Burrell argues that even transferable quotas induce a leftward shift in the supply curve. Farmers, she argues, bid up the price of quotas such that the returns are comparable to other enterprises, causing an upward displacement in the medium-term supply curve. This is true, but the capitalization of rents would also occur with the other instruments evaluated in our paper. Rents in the two-price plan would have to be allocated with marketing quotas or certificates of some kind, generating a value just like production quotas. An open-ended price support scheme with or without a co-responsibility levy will result in the values of other fixed assets like land bid up so that the returns to farming will be comparable with alternative uses of capital. If land purchases are undertaken with borrowed capital, then the same amount of "creaming" of income support by financial institutions will occur. There is nothing unique to quotas in this entire discussion by Burrell.⁵

A further criticism of our paper made by Burrell is that aggregate producer welfare is not the appropriate political target. We acknowledge that an improved and more refined policy performance measure other than aggregate producer surplus is desirable. Burrell argues that the true political target is returns to farm-owned assets of marginal full-time farmers or producer surplus per unit labor in agriculture.⁶ She states that EC agriculture suffers from too many farmers of inadequate size and that politicians ignore costs of adjustment. Indeed, Burrell implies that the price support schemes of the CAP are intended to facilitate structural rationalization of European agriculture. We question the views taken by Burrell on this issue. Indeed, we would argue the opposite in that the political rationale of price support schemes is to compensate farmers for the costs of adjustment and to slow the migration of farmers out of farming and rural areas in general. Politicians, therefore, neither ignore adjustment costs nor feel there are too many farmers. Surely, if structural adjustment was the primary motivation for the CAP, a more efficient instrument than open-ended price supports could have

been found. Quotas have the further advantage of allowing governments to achieve political efficiency more easily by being able to target benefits of price supports to specific farmers. Ironically, Burrell finds it paradoxical that the structural weakness of European agriculture is both a product of the CAP and of political constraints. This would not be a paradox if one understood the true political goals of the EC.⁷

Burrell also questions our discussion on the terms-of-trade effects of these policies and asks whether the endogeneity of supply conditions with respect to EC policy in other trading countries is incorporated in our analysis. We borrowed the excess demand elasticities from the work of Tyers and Anderson. These elasticities were estimated by using information on the domestic supply and demand elasticities in the most important trading nations, the shares of production and consumption that are traded, and the price transmission elasticities. The latter include the effects of policy in other countries by summarizing how domestic prices (and policy instruments) set by governments respond to changes in world market conditions. In this sense adjustments by other countries to EC policy is incorporated in our analysis by these price transmission elasticities. We do, however, question Burrell's assertion that the rest of the world curbs export capacity in response to any deterioration of the terms of trade resulting from EC export surplus disposal schemes. The U.S. response to increased EC export restitutions on wheat in the 1980s was to lower the loan rate and institute the export enhancement program. This U.S. response lowers the world price rather than increases it as argued by Burrell.

Finally, Burrell makes two practical points. The first point is that our welfare measure is incorrect because, under a quota system, cereal farmers with livestock would value their own feed at cost rather than at the high price support. This is absolutely false. Our measure of producer welfare is an economic one that includes opportunity costs. Selling grain to your own livestock enterprise does not preclude the economic benefits of the price support from being captured. Any farmer that feeds his own feed under a quota scheme for cereals incurs the opportunity cost of not selling the cereal on the market at the high price support. The second practical point raised involves the administrative costs of enforcing quotas with so many farms. This may be a legitimate point, but we have two qualifications. First, if it is so complex, how is it that the Canadian Wheat Board manages marketing quotas and how does the United States

⁵ Footnote 2 in Burrell states that the price of quota values for milk in the EC reflect high marginal net returns to expansion due to economies of size. This is purely speculative. All farms receiving quotas in 1984 were required to reduce production from their historical levels. All farms therefore experienced marginal returns higher than marginal costs. Just because some farms have lower marginal costs and are able to purchase quotas in expanding output does not imply a change in the economies of size. Indeed, how can one tell that the largest farms have the lowest marginal costs at the initial level of quota? Perhaps smaller than the largest farms bought quota, and so fewer large farms now exist. Besides, one can show that positive quota values exist in an industry exhibiting constant returns to scale.

⁶ Quotas are a farm-owned asset and so returns are realized by those farmers that politicians have targeted in issuing quotas to them in the first place.

⁷ Burrell argues that more farmers would remain in farming with transferable production quotas compared to other instruments attaining the same level of producer surplus. This is not correct, as quotas typically result in lower levels of output relative to other options and transferable quotas ensure production will occur by the most efficient number and size of firms. Therefore, it is quite possible that the number of farms declines more with production quotas relative to other policy instruments.

administer base acreage? Second, it is not obvious that quotas are any more difficult to administer than, say, a two-price plan or a co-responsibility levy. Both are difficult and potentially messy to administer. Nevertheless, we admit that our welfare analysis, like most, ignores administration costs.

In conclusion, Burrell argues that, because we ignore endogeneity of structural change, our policy recommendations are based on a false identity between the properties of our model and the realities of CAP reform. From the discussion above, it is obvious that Burrell has failed in developing a logically consistent analytical and empirical framework to incorporate the structural impact of EC policy alternatives. Although such an analysis was beyond the scope of our original paper, we do not wish to deter the profession from analyzing such issues. Examples of work in this vein are the landmark paper by Krueger in analyzing the differential impact of import tariffs versus quotas, and de Gorter and Fisher who analyze rent seeking in a dynamic analysis of how (non-quota!) policy in the United States induces additional deadweight losses beyond those captured in the traditional static analysis.⁸

[Received October 1990; no revision.]

⁸ For an excellent survey article, see Brooks and Heijdra.

References

- Brooks, M., and B. Heijdra. "An Exploration of Rent Seeking." *Econ. Rcd.* 65(1989):32-50.
- de Gorter, H., and E. Fisher. "The Dynamic Effects of Agricultural Subsidy Programs in the United States." Dep. Agr. Econ. Work Pap. No. 89-9, Cornell University, Nov. 1989.
- de Gorter, H., and A. Schmitz. "An Economic Analysis of the Arizona Dairy Market." Preliminary Technical Paper, United Dairymen of Arizona Case, Dep. Agr. and Resour. Econ., University of California, Berkeley, Nov. 1981.
- Krueger, A. "The Political Economy of the Rent-Seeking Society." *Amer. Econ. Rev.* 64(1974):291-303.
- Public Hearing on Federal Milk Marketing Orders. Testimony of Edward V. Jesse on behalf of Upper Midwest Federal Order Coalition. Minneapolis MN, Sep. 1990.
- Rucker, R. R., W. N. Thurman, and D. A. Sumner. "Production Rights with Limited Transferability: A Case Study of U.S. Tobacco and Peanut Programs." Dep. Agr. and Resour. Econ., North Carolina State University, Aug. 1990.
- Runge, C. Ford, and H. von Witzke. "Institutional Change in the Common Agricultural Policy of the European Community." *Amer. J. Agr. Econ.* 69(1987):213-22.
- Tyers, R., and K. Anderson. "Imperfect Price Transmission and Implied Trade Elasticities in a Multicommodity World." *Elasticities in International Agricultural Trade*, ed. Colin Carter and Walter H. Gardiner. Boulder CO: Westview Press, 1988.

Books Reviewed

Atkin, Michael. *Agricultural Commodity Markets: A Guide to Futures Trading*. London: Routledge, 1989, 249 pp., \$79.95.

Until the late 1970s, agricultural economists had a virtual monopoly on books dealing with the futures markets. With the development of financial and options contracts, a host of books were generated by authors who are not agricultural economists, although a recent book by Leuthold et al., *The Theory and Practice of Futures Markets*, bucks the trend. Many books include agricultural commodities as only one segment of the futures market complex, for example *Understanding Futures Markets*, by Kolb. Other books focus on specific nonagricultural futures contracts, the Loosigian books, *Interest Rate Futures* and *Foreign Exchange Futures*, are good examples. The current approach is not to examine specific types of futures contracts at all, but to present theory and analytical techniques useful to advanced students and practitioners. Examples of this genre include *Option Pricing* by Jarrow and Rudd and *Commodities Trading* by Seidel and Ginsberg.

The book under consideration, *Agricultural Commodity Markets: A Guide to Futures Trading*, seems to be a return to a bygone era, a descriptive book focusing only on agricultural commodity futures markets. In the author's words, it is a book "aimed at a broad audience of students, economists, traders and the general public" (p. xv). The book is intended to fill the gap between highly technical publications and the "get-rich-quick trading commodities" books.

Organizationally, the book is sectioned into two parts. The first part gives the reader an overview of the structure and operations of the agricultural commodities futures markets. In the second part of the book, the author examines the major agricultural commodities including grains, oilseeds, livestock and meat, coffee, cotton, sugar, and cocoa.

The book is well written, although the use of British rather than American spelling in some places is slightly disconcerting. More disturbing is the use of British terms, such as contango for carrying charge, backwardation for inverse carrying charge, highly geared for leveraged and maize for corn. The author does explain the difference in terminology. The first part of the book on the structure and operations of futures markets is rather brief, although the chapter on options is especially well written. The discussion of major agricultural commodities is the usual description of factors affecting supply and demand conditions for each commodity. The treatment of hedging takes only slightly more than three pages, although it is mentioned in passing in several other places. In contrast, an entire chapter is devoted to speculation. A major drawback in both chapters, throughout the book for that matter, is the lack of examples.

Atkin's last chapter, "The Future of Futures," is

the best in the book. He traces the development of futures markets, including the introduction of financial futures contracts, to the present time and then describes his vision of the future for futures trading. He envisions the development of a global market for futures trading and points out the need for futures trading centers rivaling Chicago to be located in the Pacific Rim and in Europe. Atkin also points out that "the potential for increased developing country usage of futures markets is huge, given their importance in world agricultural trade" (p. 222). Intriguing is his discussion of the development of computer-based, electronic trading systems which would allow 24-hour trading around the world. Because of the difficulties which would arise if live floor trading were replaced with totally computerized markets, Atkin foresees "a novel combination of the two types of trading . . ." (p. 216).

In summary, *Agricultural Commodity Markets* would not make a good teaching text for the simple reason that it lacks examples, problems, and end-of-chapter questions that would aid student understanding. Most economists today would expect more depth and breadth in a book on the futures markets. Beginning traders and the general public interested only in agricultural futures would find the book useful, although potential hedgers would be very much disappointed.

Michael W. Woolverton
American Graduate School of
International Management

References

- Jarrow, Robert A., and Andrew Rudd. *Option Pricing*. Homewood IL: Richard D. Irwin, 1983.
Kolb, Robert W. *Understanding Futures Markets*. Glenview IL: Scott, Foresman and Company, 1985.
Leuthold, Raymond M., Joan Junkus, and Jean Cordier. *The Theory and Practice of Futures Markets*. Lexington MA: Lexington Books, 1989.
Loosigian, Allan M. *Foreign Exchange Futures*. Homewood IL: Dow Jones-Irwin, 1981.
———. *Interest Rate Futures*. Homewood IL: Dow Jones-Irwin, 1980.
Seidel, Andrew D., and Philip M. Ginsberg. *Commodities Trading: Foundations, Analysis, and Operations*. Englewood Cliffs NJ: Prentice-Hall, 1983.

Becker, Tilman. *Die Weizenexportpolitik der Europäischen Gemeinschaft* (Wheat Export Policy of the European Community). Agrarwirtschaft Special Study No. 121. Frankfurt am Main, Federal Republic of Germany: Alfred

Strothe Publishers, 1989, 242 pp., 44.5 DM (approx. \$27.81).

Since the Food Security Act of 1985, the United States has pursued a two-pronged agricultural export policy with respect to the European Community (EC). It has employed the Export Enhancement Program (EEP) to increase U.S. market share in EC-contested markets while encouraging multilateral trade liberalization through the General Agreement on Tariffs and Trade (GATT). Because the focus of the competition between the United States and the EC has been the world wheat market, European studies of this market are of obvious interest to an American audience.

Becker's study has two objectives: to analyze the EC grain export policy and to suggest decision criteria for efficient export administration. These objectives are pursued in five chapters: introduction, the EC as grain exporter, the world wheat market, the EC wheat export policy, and summary.

From Becker's perspective, EC grain policy faces a conflict between the stable domestic demand and rising production. EC grain yields rose at a rate of 3.3% annually between 1967 and 1984 and are expected to continue increasing until by the year 2000 the EC-12 may have to retire up to 30% of its current production area.

Because the EC grain programs are likely to remain expensive, how can export subsidy (restitution) expenses be minimized? Becker analyzes EC restitution costs for wheat and barley as an example. Using a Lagrangian framework, Becker concludes that EC expenses are minimized when the ratio of marginal changes in wheat and barley restitution costs are equal to the ratio of wheat and barley feeding values. Becker reports that the EC appears to have used this decision criterion in 1983 and 1984 but deviated from it in 1985 to avoid confrontation with the United States.

He goes on to characterize the world wheat market as oligopolistic, with the United States providing price leadership. Grain storage is the key to his understanding of the wheat market because it allows exporters to practice seasonal price discrimination. By manipulating stocks and export subsidies, exporters can force the elasticity of export demand toward 1 and maximize export earnings. U.S. price support policy reinforces this result, Becker argues, when target prices are set in the elastic portion and loan rates are set in the inelastic portion of the demand curve.

The author presents a graphical analysis of supply and demand conditions in the subsidized and non-subsidized export markets and concludes that export subsidies are rational only when revenue gains from price increases in nonsubsidized markets offset costs. This occurs when the supply curve in the nonsubsidized market is inelastic. Consequently, unless world supplies are constrained, export subsidies are likely to result in a net welfare loss.

Becker concludes by addressing two questions: when and where should the EC export?

When? Becker simulates seasonal EC export levels using a profit function for wheat storage. Studying both budget cost minimization and profit maximization, optimal marketings are shown to occur earlier in the marketing season as storage costs, interest costs, wheat prices, and the elasticity of demand rise. Becker concludes that EC welfare gains could increase 10 million ECU by using export licenses to select the optimal timing of export marketings. The same result is obtained when private firms time marketings according to his framework. Optimal timing of wheat export licenses in 1987/88 could reportedly have saved the EC up to 100 million ECU. Becker notes that cost minimization and profit maximization yield differing seasonal export recommendations.

Where? In analyzing the relative performance of exporters in the nine EC export marketing zones, Becker develops a probabilistic model of trade flows based on world origin and destination statistics. He argues that the EC should concentrate on import markets in which the competition from other exporters is unstable; that is, in markets where importers are price sensitive rather than exhibiting loyalty to one supplier. EC subsidy expenses should furthermore rise in markets with high elasticity of demand and fall in markets with low elasticity of demand. In this manner, marginal changes in restitution costs are equalized across all importers and budget expenses are minimized. This has the effect of increasing restitutions in distant markets and decreasing restitutions in local markets.

Becker more than adequately accomplishes the objectives set forth for his work. His analysis of EC export operations and policy is thorough, and his outline of decision criteria for program administrators is convincing. The study's chief weakness stems, in part, from its pragmatic nature. Becker's research is preoccupied with "second-best" policy prescriptions—prescriptions that increase the efficiency of existing programs. He recognizes that export subsidies are suboptimal in the current market environment and that land retirement cannot ultimately be avoided. He does not, however, employ these insights to widen the scope of this research. Consequently, it is clear that his research results will be of most interest to individuals who are administrators of EC wheat programs. Advocates of trade liberalization are less likely to find his research useful.

Stephen W. Hiemstra
Farm Credit Administration

Coombs, H. C. *The Return of Scarcity: Strategies for an Economic Future*. Cambridge, UK: Cambridge University Press, 1990, x + 171 pp., \$39.50, \$17.95 paper.

This book is discomfiting. Its essence is that population growth, technological change, and free markets confronting finite material resources are leading the world on the path to perdition. Of course, its theme is not new, nor are its arguments and evidence par-

ticularly original. But it packages the message in such a way that even the most ardent worshiper of free markets has his/her faith shaken a little by the end of the book.

The author has had a distinguished career in academia, government, and community service in Australia. He is a former governor of the Reserve Bank and chancellor of the Australian National University. For the past twenty years, he has been a leader in developing policies to promote the arts and policies to assist Australia's aboriginal peoples. Now, at eighty-four years old, he is a visiting fellow of the Centre for Resource and Environmental Studies at the Australian National University.

This book is a compilation of essays written over the past two decades. The essays reflect the author's "anxiety as a private citizen and member of the academic community about the threat to our society" arising from the interaction of economic and ecological problems. As might be expected, there is considerable repetition of ideas in the nine essays. Also, there is consistency in his essays, which is somewhat surprising with worldwide changes in political and economic systems over the past twenty years.

According to Coombs, market economies produce a vast array of goods, and by standard measures, GNP and accumulated wealth, western societies' economic welfare is increasing. But the end products of the system—the lifestyles of its members, their cultural experiences, their personal relationships, and their physical environment—are worsening and are likely to continue to do so.

Population growth, increased demand for consumer goods, increased scarcity of basic resources, increased use of labor-saving technology, and the emergence of a world economy continually change the operation of the economic system. First, they increase rents for owners of scarce resources, owners of technology, or those with specialized knowledge. Labor is rewarded less, and it is placed in an increasingly unfavorable bargaining position. The result is a continuing redistribution of income toward owners and away from workers and providers of services. Increasing shares of income go to transnational corporations, which tend to dominate the development of economies. There is increasing difficulty in satisfying the expectations of adequate and improving real incomes for most in society. Expanding population, technological innovations, and environmental degradation reduce practical alternatives for a simpler lifestyle. Private property closes off opportunities for low-cost access to the amenities of nature. Environment quality deteriorates as technological change results in polluted air and water. Species are threatened with their very survival. Quality of life may improve for the few fortunate, but for the vast majority, quality of life deteriorates even though their consumption of consumer goods may increase.

Coombs solutions mostly rely on heavy handed government intervention. Some of this intervention would change economic incentives. A resource tax

or graduated royalties would be levied against users of scarce resources. Differential pricing for electricity, petroleum, and water would result in higher prices for heavier users. A tax rebate would be used to encourage savings and investment. Advertising would be thwarted by limiting income tax deductibility of advertising expenses. But his principle strategies are selected government ownership or regulation. Scarce minerals, forests, seas, soils, vegetation, and wildlife would be collectively owned. Title to these scarce resources would be vested in an authority independent of corporate and political control. License for their use would be granted to those assuring their conservation. Monopoly rents would be collected and distributed to subsidize resource conservation research, to finance community facilities designed to improve the quality of life, and to provide all citizens a "national dividend." Land users including farmers would submit land use management plans and have them approved by a conservation authority. Similarly, all enterprises involved in resource industries would comply with approved resource and environmental plans. The government would use unemployed labor to protect and improve the environment. Inheritance taxes would confiscate most decedents' property except the family home, personal possessions, family-operated farms, and other individually operated enterprises.

Those who see the "workings of the market as expression of divine wisdom" will be irritated by this book. Indeed, they probably will dismiss it after a few pages. After all, resource substitution, technological change, and increasing productivity will mitigate the impacts of increasingly scarce resources. Societies do not need the heavy hand of government to adapt. In fact, market mechanisms provide for quicker and more efficient adaptation to resource scarcity than do administrative solutions. I imagine that Coombs would concur with this judgment. He does not question the efficiency of the market or its ability to respond to scarcity. His argument is that standard economic measures of performance are faulty. They do not value quality of life—income distribution, environmental quality, stocks of finite mineral resources, amenities of nature, opportunities to choose simpler lifestyles, the involvement of individuals in groups, and the fulfillment individuals receive in their lives. His view is that quality of life is deteriorating even though GNP is increasing and that only collective action to cause structural change in our economic system can keep this deterioration from accelerating.

D. Lynn Forster
Ohio State University

Demissie, Ejiguu. *Small-Scale Agriculture in America: Race, Economics, and the Future*, 1st ed. Boulder CO: Westview Press, 1990, x + 135 pp., \$20.00.

This book accomplishes what few others, if any, have attempted, an economic assessment of small-scale

farms (gross sales less than \$40,000). This assessment is vital because small-scale farms represent the majority of agricultural enterprises in the United States. Moreover, the author provides an analysis of both black and white small-scale farmers. The primary objective of the book is to describe the history and current status of small-scale farm operators in the United States and to provide background information that can be used by policy makers and program administrators as they attempt to halt further decline of small-scale farms. Demissie notes that the book is written for audiences in the agricultural and nonagricultural communities, students of agriculture, as well as citizens who have concerns about small-scale agriculture and rural development.

The book is divided into seven chapters. The introductory chapter begins with the author's description of the structural changes (farm size and number) that have occurred in the agricultural sector this century. This chapter clearly shows that "most of the decline in the number of farms has been in the small-scale farm category" (p. 3).

In chapter 2, the general characteristics and trends for small-scale farms and farm operators are addressed. The author provides a "profile of small-scale farms with respect to their agricultural production, resource use and control, farm asset holding, types of enterprises, and farm production expenses" (p. 8). Furthermore, he compares and contrasts the profiles over time and by farm size, thus showing the rapid decline of farms, especially among blacks.

Chapter 3, entitled "Characteristics and Composition of Black Farm Operators and Their Farms," examines some of the same characteristics addressed in chapter 2 but discusses differences between black and white farm operators. It presents an historical view of these farms as well as their current status. The chapter concludes with a short discussion of the historical development of 1890 land-grant institutions (black land-grant colleges and universities) and their role in assisting small-scale farmers, especially black farmers.

Chapter 4 gives an economic assessment of the farm and farm operators by examining their past and present financial conditions. The chapter looks at trends in farm income, assets and debts, surmising that their financial problems are symptoms of the structural changes. The financial comparison was based on farm-sales class, thus depicting the poor financial condition of small-scale farmers. This chapter further showed that many small-scale farmers do not obtain financing from traditional lending institutions but from internal savings.

Chapter 5 discusses some of the social, economic, and political constraints that have negatively impacted the survival of small-scale farm operation in the United States. The factors addressed included resource endowment, education, technology, marketing, management and information, government policies and programs, and sociological constraints.

In chapter 6, "Small-Scale Agricultural and Rural

Development," the author attempts to show that the loss of small-scale farms has had a substantial impact on both the viability of rural communities as well as the fabric of rural life. Although this discussion is interesting, it lacks depth. Very few data are used to support the discussion.

Chapter 7 provides a brief discussion about the various programs (technology, management, marketing, education, credit, and government) that can be useful in solving problems faced by small-scale farmers. The author places particular emphasis on the enhancement of research and extension efforts to eliminate the decline of small-scale farms.

The material contained in this book, basically taken from the *Census of Agriculture*, can contribute to an enlightened dialogue about the plight of small-scale farmers, especially black operators. It is an excellent resource book that can be used to increase public awareness of the issues, concerns, and possible public policies that could affect the future survival of small-scale farmers. Moreover, although much of this material is not new, having the information in this condensed version makes it valuable. The book falls short in several areas, however. It is not well integrated, there are some redundant materials and poor sentence structure, especially in the first two chapters. Because it is written for the average citizen, the book may not have the sophistication that graduate students and professionals in the field expect. Last, the future, as depicted in the book's title is not adequately addressed. Notwithstanding, the book makes a substantial contribution to a very important segment of our population, small-scale farmers.

Donald R. McDowell
North Carolina A&T University

Ferguson, Roy C., II. *Managing for Profit in Commercial Agriculture*. Englewood Cliffs NJ: Prentice-Hall, 1990, xvii + 208 pp., \$47.20.

This book attempts to deal with both the Canadian and U.S. situation, but the Canadian references are an obvious afterthought. The book has several weaknesses but is useful for farm management professionals and students as a reference source. The first and most glaring weakness is that the task of proper analysis of the farm business is made to look too complicated for most farm business managers. The calculations and the acronyms used in the calculations are often confusing. The author has made little attempt to clarify and simplify the text. The amount of detail recommended leaves the impression that many managers will need to hire professional expertise to help manage their businesses.

The book begins by discussing the farm business problems of the past ten to fifteen years and the role that weak financial accounting and analysis procedures played in these problems. The author argues that most of these problems could have been avoided by following his system of financial analysis. He correctly stresses that agriculture is no different from any

other form of business and therefore should be subject to the same strict borrowing, accounting, and analysis rules as other businesses. He also argues that detailed accounting information with as much accuracy as possible should be emphasized. In chapter 4, he presents the recommended financial documentation needed for agriculture and includes examples of balance sheets, operating statements, cash flow statements, and loan documentation requirements. The amount of detail recommended is very commendable but probably impossible to achieve economically on many Canadian farms, even if the owners wanted to.

Chapters 5, 6, and 7 present the "Ferguson System of Financial Analysis," and chapter 8 includes a number of cases that apply the system. Fourteen "signals of financial health" are presented in chapter 5, with emphasis that this is the minimum for proper financial analysis and that the reduction to fourteen already represents a significant compromise. The full "Ferguson System" of thirty-six financial ratios and indices are presented in chapter 7. The author emphasizes that no single ratio or small group of ratios should be used on their own or with less than four-to-six years of detailed and accurate data to back them up. These data requirements both in duration and detail effectively eliminate most Canadian farms.

The author goes on to designate a series of critical values or guidelines for the various ratios and indices dependent on enterprise type, such as crops, hogs, cow/calf, beef feedlot, poultry, and dairy. These guidelines are given as precise measurements but are backed up only by the "Ferguson Group experience" rather than with any quantitative study proving them to be legitimate leading indicators of farm financial troubles. In addition, the ratios should be adjusted by a rather complicated process for changes in such items as percentage of total revenue and proportionate sales from the various enterprise types referred to above. Chapter 6 presents examples of how to calculate these adjusted guideline levels. Chapter 8 presents a number of cases to illustrate the use of the "Ferguson System." The cases are quite detailed, and the text describing the analysis is hard to follow. I could not help but get the impression that with detail supplied in the cases, the answers for the most part were quite clear by studying the exhibits and did not need the full application of the "Ferguson System" to solve. In fact the "Ferguson System" application seemed to add to the confusion.

Chapter 9 is entitled "Eight Fatal Flaws Leading to Problem Debt" and is a very good summary of some important considerations in farm credit management. Chapter 10 deals with farmland purchase economics. The return to land are calculated for a number of crops, including grassland, land quality, and locations. The author concludes that land seldom if ever pays for itself and that farm managers should seriously consider renting a portion of their land.

Chapter 11 is entitled "Capital Development" and deals with the subject of obtaining capital from lenders and investors. The author presents a number of

hints that will ensure the lender's thoughtful consideration of the loan proposal. The chapter also deals with ways and means of attracting investor capital and suggests the completion of a 200- to 400-page complete business plan. I am not aware of too many farmers who could complete such a plan without considerable professional help.

Chapter 12 deals with the understanding and effective implementation of five disciplinary principles called "Management Stardom's 5 Points." The first is knowing the business you are in, which includes the key characteristics and requirements of commercial agriculture and the need for a strong management information system. Next is the requirement to maintain a correct internal business structure with respect to the use of capital and monitoring performance. Third, eliminate price risk for both inputs and outputs through futures markets and forward contracts. The fourth point is to maintain efficient operations by reducing cost, eliminating crisis management, and monitoring performance. The final point is the use of effective people from management and ownership to employees to professional counsel.

The final chapter is an epilogue that warns that the recent financial disasters could easily be repeated if farmers do not become significantly more business conscious and adopt the approach presented in the book.

In summary, the book is a good addition to the library of anyone interested in farm business management. The frequent references to the author's system and company and the need for professional help eliminate it, in my opinion, from use as a university text.

William J. Brown
University of Saskatchewan

Hancock, Graham. *Lords of Poverty: The Power, Prestige, and Corruption of the International Aid Business*. New York: The Atlantic Monthly Press, 1989, 234 pp., \$17.95.

Western nations provide some \$60 billion of foreign aid annually to developing or Third World countries. Despite this huge transfer of funds from the West, Hancock argues that developing countries would be better off without it. The author's assessment of foreign aid is based on his many years of foreign experience as a reporter for *The Economist*. In his words, foreign "aid is not bad . . . because it is sometimes misused, corrupt or crass; rather, it is inherently bad, bad to the bone, and utterly beyond reform" (p. 183). To illustrate the ill effects of foreign aid, Hancock attacks the "rich and powerful bureaucracies . . . that administer the West's aid and then deliver it to the poor of the Third World . . ." (p. xv).

Hancock's book is divided into six chapters, with each chapter attacking specific aid bureaucracies or employees of the aid industry. Chapter 1, titled "Masters of Disasters," sets the stage, and it provides

many examples of development gone awry. Included among these are examples of charitable organizations spending most monies intended for developing countries, food shipments never reaching their intended destinations, donated food being unfit for human or animal consumption, field employees of the aid industry refusing to do field or site assessments of food and medicine needs during disasters or crises, and western experts placing evangelism ahead of assistance. Chapter 2 attacks the World Bank, the International Monetary Fund, and other multilateral organizations such as the European Economic Community and the United Nations. The alleged sins of these organizations include wasteful expenditures of \$10 million for annual meetings of the World Bank and International Monetary fund; structural adjustment loans (SAL) causing massive increases in unemployment in Chile, substantial increases in child malnutrition in Peru, and significant increases in infant mortality rates in Brazil.

Chapter 3, "The Aristocracy of Mercy," attacks the many organizations of the United Nations and the many individuals who profit from these and other aid-related organizations. Hancock states that most workers in the aid industry admit they are attracted to the high salaries and the opportunity to stash away small fortunes. An example includes a UN worker who states, "I won't try to pretend that I took this job out of idealism . . . the money was very important to me." (p. 79). The monetary rewards of "aid industry" employment also are evident as one considers the fact that another employee of the United Nations privileged aristocracy "was receiving a subsidy of \$10,661 per month towards the cost of his three-bedroom apartment . . . close to UN headquarters" (p. 93). Moreover, the Food and Agriculture Organization arm of the United Nations has "750 individuals whose pensionable remuneration ranges from \$70,000 to \$120,000 a year . . ." (p. 93). Hancock argues that such remunerations are not paid because they are necessary to attract the best and most competent employees. On the contrary, "at every level of the multilateral agencies, maladjusted, inadequate, incompetent individuals are to be found clinging tenaciously to highly paid jobs . . . betraying the world's poor in whose name they have been appointed" (p. 99).

Chapter 4, "The Midas Touch," attacks foreign "experts" for planning and implementing projects in the Third World on the basis of information gained from runways and highways. Hancock states that "major development projects absorbing vast sums of Western tax-payers' money and affecting the lives of billions of people in the Third World for years to come are routinely conceived, planned, supervised and appraised on the basis of visits that are equally brief, hurried and superficial" (p. 121). He concludes that such projects occur because the World Bank and its partners in development are "out of touch with poor people, and with the day-to-day realities of their lives . . ." (p. 151).

Hancock's attack in chapter 5 is directed at the dis-

proportionate share of "aid" expenditures that are spent in industrialized countries. He questions the logic of an American aid program that is intended "to meet the basic needs of poor people in developing countries" (p. 160) . . . when "70 cents out of every dollar of American assistance to the Third World never actually leave the United States" (p. 156). Similarly, he questions the Third World development efforts of the World Bank when 70% of its funds are "spent on goods and services from the rich industrialized countries" (p. 159). Worse still, purchased goods and services from industrialized nations are often marked-up and of low quality. He further argues that the introduction of food commodities, such as wheat, does a disservice to the development process because it relegates local commodities, such as corn, to an inferior status.

The final chapter summarizes the previous five but also provides additional evidence about the ill effects of foreign aid. A point of focus for the author relates to financial flows between Third World countries and wealthy industrialized nations. Measures of financial flows show that Third World nations are net donors to industrialized nations, raising the question of who is helping whom?

Any reader of this book, whether employed inside or outside the aid industry, will probably agree that the author provides evidence that there have been many ill-conceived, mismanaged, and wasteful projects undertaken in Third World countries. Individuals inside the aid industry, and even some outside, will undoubtedly conclude that the author has provided a biased and unfair assessment of the aid industry, however. Despite the billions of dollars that have been wasted, clearly there are some success stories related to foreign aid. Improved education, better transportation, telephone communication, and electricity generation are benefits that most observers attribute to foreign aid. Perhaps net financial flows between the rich and poor are not a relevant measure of foreign assistance. A more meaningful measure might be the enhanced economic ability of poor countries to make financial transfers as a result of previous foreign assistance.

Additionally, perhaps foreign assistance should not be evaluated by the extent to which poor people are involved in project planning but the extent to which they benefit from project outcomes. In this regard, economic development does seem suspect. Even in the United States where poverty is less pressing compared with the Third World, evidence suggests that the poor people today are worse off than they were ten years ago. Indeed, there are individuals who readily admit that the war on poverty and other aid programs to help the poor have been a failure in the United States. If poor people are made worse off by poverty programs in industrialized nations where the problems are less severe, then one has to wonder how the aid industry can deal with these problems in the Third World where problems are more severe.

This book illustrates that it is easier to write about

the problems of poverty than it is to solve them. Nevertheless, the book provides an assessment of foreign aid that is sure to challenge the actions and thought patterns of aid workers. It is an inexpensive book that can be purchased with a fraction of one's hardship allowance.

Eugene Jones
Ohio State University

Harl, Neil E. *The Farm Debt Crisis of the 1980s*.

Ames: Iowa State University Press, 1990, xix + 305 pp., \$24.95.

This book is part of the Henry A. Wallace Series on Agricultural History and Rural Studies, a series designed to explore the many aspects of agricultural and rural life within historical perspectives. The author was prompted to write the book "by a strong belief that the lessons learned in the farm debt crisis of the 1980s should be identified and memorialized to assure ready access on the occasion of the next major economic downturn in agriculture" (p. 269). The book thus represents a history of one participant's personal odyssey through the farm debt crisis period of the past decade. The book was conceived as history, not economics, and should, in fairness, be judged as such. Harl's intended audience is future policy makers, who will learn his view of appropriate intervention in the event of future such crises.

The author portrays himself as a visionary whose role was to arouse the nation to recognize what he feels was the irresponsibility of the Economic Recovery Tax Act of 1981 and in his opinion the tragedy it produced in the agricultural sector. He champions public sector intervention to address this tragedy since, in his view, it was a result of bad public policy foisted on an innocent and powerless agricultural sector. Thus, beyond a broad synopsis of the author's opinion as to the causes of the crisis and the general innocence of farmers and (less so) their lenders for the calamity they faced, the major themes of the book are the need for massive public sector intervention and the recalcitrance of federal-level public officials in acknowledging and addressing this need. In particular, the intervention Harl advocated was widespread farm debt restructuring and warehousing of farmland acquired because of bad loans in a federally chartered Agricultural Credit Corporation (ACC) until times improved and farmland values recovered.

The volume is divided into ten chapters dealing with different aspects of the debt crisis. Chapter 1 describes the public policies that Harl feels inexorably led to the crisis, while chapter 2 presents his evidence on the innocence of farmers and the slowness with which all concerned realized the magnitude of their plight. Chapter 3 describes the need for and chronicles Harl's proposals for debt restructuring. Chapter 4 relates the growing problems of the Farm Credit System (FCS) and the author's attempts to convince some of the FCS leadership that their system was vul-

nerable to crippling loan quality problems. Chapters 5 and 6 address proposals for further federal involvement, heel dragging in Washington, and "the obvious gap between reality and what Washington seemed to believe was reality" (p. 164). Chapter 7 lauds state efforts to address the farm debt crisis. Chapter 8 discusses the impacts of the farm debt crisis on local businesses, and chapter 9 addresses at length the influence of the media during the farm debt crisis. In keeping with Harl's purpose for writing this book alluded to earlier, the final chapter (10) summarizes his twelve lessons for posterity. However, the author makes no attempt to develop the relative importance of or logic behind these lessons, nor do all of them directly follow from preceding chapters.

Of most interest to academic readers will be chapters 1, 2, and 10 as a source of testable hypotheses for future scholarly research. Portions of chapters 3 through 8 will hold interest for some readers concerned with their particular focuses, although they will pay the price of wading through a chronologically disorganized and ponderous account. Harl relies heavily on references to or quotations from the popular press and trade publications, many of which add only peripheral detail and tedium.

In the initial chapters, Harl presents his view that the farm debt crisis arose as the result of three federal policies: (a) the policies pursued over at least five successive administrations with respect to inflation, (b) the abrupt action taken by the Federal Reserve Board beginning in October 1979 to reduce the rate of inflation, and (c) the massive experiment implicit in the large tax cuts of 1981. He holds these policies directly responsible for a strong dollar, lower exports, high interest rates, falling land values, and sizable farm loan defaults. In contrast, the author repeatedly argues that farmers affected by debt problems during the decade were neither greedy, nor poor managers, nor speculators. Despite the rapid build-up in farm debt in the 1970s, he is grudging in assessing farmers much responsibility. The most he admits is that a substantial portion of the farmers in deep financial difficulty were striving to reach the minimum point on the average cost curve "in perhaps less time than would have been prudent, at least viewed with 20/20 hindsight" (p. 70). Because the federal government is viewed as initiating the forces leading to the severe farm sector debt trauma and farmers were generally innocent, Harl suggests that such assistance was economically justified whenever the total costs and consequences of intervention were less than the expected costs and consequences of failure to intervene. He never explicitly identifies these costs, quantifies them, or discusses on what basis they should be distributed among lenders, taxpayers, and borrowers.

The book fails to address the debt situation in U.S. agriculture as it developed in the 1970s and to analyze the sector's economic health as it went into the 1980s. During the 1970-80 period, at the same time the GNP deflator increased 104.0%, U.S. per acre

farmland values grew 277.9% (369.4% in Iowa). Total farm debt expanded 228.8% (291.4% through 1983) despite considerable excess capacity and reliance on federal subsidies. Even Farmers Home Administration lending grew 504.6% during the 1970s. This lender of last resort loaned \$11.9 billion (\$19.2 billion for 1970–82) alone under various types of emergency loan programs during a decade that was deemed a prosperous one for agriculture.

Harl recognizes that the problem during the 1980s was too much farm debt concentrated among a subset of farmers. He then recommends some expensive solutions. He estimated that his debt restructuring proposal would cost the federal government \$20 to \$30 billion if adopted, and the tab for the ACC would have been up to \$10 billion (pp. 74, 152). The reviewers feel one problem with the ACC concept is that stabilization programs need to be in place *ex ante*; they do not work *ex post*. Nor do they work if the asset prices have moved to a lower equilibrium based on economic fundamentals rather than overshooting. By the time Harl proposed the ACC in late November 1984—early 1985, U.S. farmland values already had declined about half of the amount they ultimately fell (in Iowa about 45% of the ultimate 60.7% slide had already happened).

A major shortcoming of the book is its questionable perspective and lack of objective analysis. In this vein chapter 7, "States to the Rescue," deals with state involvement in the farm debt crisis. This is misleading, juxtaposed after two contentious chapters indicating federal inaction. State credit program subsidies, while significant, are small compared to those associated with federal farm programs. For example, in 1987 and 1988, state credit subsidies were roughly 2% of federal credit program subsidies to farmers. There is a neglect of the linkages among the major federal agricultural programs throughout the book and the fact that the federal support to agriculture during the 1980s was massive. Net CCC outlays to farmers, for instance, during the 1980–90 period totaled \$140.3 billion. Direct government payments alone to Iowa farmers during 1980–89 totaled \$8.47 billion, or about \$3,000 for every Iowan, or \$81,500 for every Iowa farm existing at the end of the decade. Federal policy makers felt, given the massive level of support, that new large specialized credit programs, such as the ACC plan, were too expensive and created perverse incentives for farm decision makers. Federal assistance tended to take the form of income rather than credit programs, hence helping all farmers and avoiding the equity issues of who should and should not receive help.

This book is a parochial and often strident memoir by a major player in the farm debt policy debates of the 1980s. While it may be of interest to some readers for that reason, serious scholars of the crisis will not find this to be an analytical or balanced view of the 1980s farm debt crisis. There remains a need for a workmanlike, scholarly analysis of the period including the economic status of the sector prior to the

1980s, the causal factors of the crisis, the effectiveness of and interrelationships among credit and other policies, and the status of the farm sector by the end of the decade.

Jerome M. Stam and
Robert N. Collender
Economic Research Service
U.S. Department of Agriculture

Morse, George W., ed. *The Retention and Expansion of Existing Businesses: Theory and Practice in Business Visitation Programs*. Ames: Iowa State University Press, 1990, xvii + 195 pp., \$21.95.

Morse and six collaborators wrote this short book which describes the theoretical foundations, educational programs, and practical applications of the business retention and expansion visitation program (R&E), a local economic development technique which was first tried in the 1960s. The book is designed to serve as a basic reference for state development officials, local practitioners of local economic development, cooperative extension, and community economics academicians wishing to start state-level educational and technical assistance programs. According to the authors, their goal "is to stimulate local economic development by helping existing businesses increase their sales and employment through a combination of programs that not only increase these firms' competitiveness but also improve the community's quality of life."

R&E programs have gained considerable popularity in the 1980s. Unlike the informal, unstructured predecessor programs used by communities in the 1960s and 1970s, current R&E programs are more sophisticated and structured. They blend applied research with public relations and focus on "immediate action for resolving individual firm problems and long-term improving the communities' overall business climate and the competitiveness of local firms." This book provides the first in-depth analysis of these new, highly structured R&E visitation programs. Most of them operate out of land grant universities, but a few are operated by state governments and utilities.

The authors outline a multiple goal approach for R&E programs based upon their experience and research primarily in the state of Ohio, where these programs have been used most extensively. Part I, *The Theoretical Foundations of R&E Visitation Programs*, presents a conceptual model for the R&E visitation program and describes the impact of existing firms on regional growth. In chapter 1, Morse defines R&E visitation programs and provides a conceptual model. He argues that R&E visitation is more than a public relations program, more than an attempt to help firms move out of the past, and more than improved firm efficiency. In chapter 2, Miller provides empirical evidence contrasting the impact of existing firms on a community's economy with the

impact of new firms. His findings support earlier research that existing firms have greater impact on local economic development than new firms. Part II, Educational Programs, outlines the methods and objectives of the educational/technical assistance programs provided to communities starting new R&E visitation programs. In chapter 3, Otto, Morse, and Hagey review the status of current R&E programs and conclude that most states now recognize the importance of working with existing businesses to stimulate economic development, but few have developed training programs to facilitate the process. In chapter 4, Morse and Hagey present a detailed outline of the fifteen steps involved in conducting a local R&E visitation program based on their experience in Ohio.

In chapter 5, Morse points out the importance of supplementing primary data gathered from the visitation program with secondary data. He notes that R&E visitation data are much more meaningful "when studied in the context of national and regional economic and demographic changes, common industrial location factors, and the availability of state and federal assistance for economic development." Chapter 6, written by Rohrer and Morse, addresses the question, "How can needed information be provided efficiently and effectively to local volunteer teams?" In their opinion, this is best accomplished by presenting the Economic Development Teleconference Series to communities. The importance of the final report and some potential applications of the information gained are explained in chapter 7.

Part III of the book, Successful Local R&E Visitation Programs, outlines the structure of successful local programs in five states. Chapter 8 describes the major types of benefits in R&E programs, and chapters 9 and 10 list impacts of the R&E program in two rural counties in Ohio. Chapter 11 reports McLaughlin's research on an Ohio R&E program during its actual implementation. In chapter 12, Bentley draws on the extensive research done on plant closings and preventing layoffs by Hansen and Bentley to answer the question, "What do you do if the R&E program fails to prevent a plant closure or downsizing in the community?" Several assistance strategies are described, and the importance of having the assistance of a neutral, third-party facilitator or catalyst is stressed. Several appendices are provided that contain copies of the forms used in the Ohio R&E program.

The Retention and Expansion of Existing Businesses is a concise, well-written book which describes the process and usefulness of R&E visitation programs. Practitioners of the development craft will benefit from reading this book and applying some of the ideas it contains. This development approach emphasizes the importance of helping existing firms instead of spending time and resources trying to "steal" firms from other communities and states. Equally important, its focus on the R&E programs represents a new, forward-looking activity for land grant university extension services rather than continuing to de-

vote scarce resources to traditional but anachronistic activities. Unfortunately, many land grant universities are investing only token resources in these newer, more important programs.

Perhaps the biggest weaknesses of the R&E approach are its complex process and the substantial external training and technical assistance needed to make it successful. One suspects that few R&E programs outside Ohio measure up to the high standards set out in the book. Heavy dependence on volunteers and a lack of training and close supervision often lead to superficial studies and little follow-up.

The usefulness and application of this book can be strengthened by exploring the relationships and linkages which can and should be developed by university extension or other R&E staff with other players in the economic development arena to ensure concrete R&E results beyond a final banquet or community meeting. In some states it appears that R&E efforts are being carried out in isolation and are not fully integrated with other development agencies and groups in the community or even with other units on a university campus. Equally serious, R&E programs may be treated as a fad or a one-shot affair, soon to be discarded as development or extension administrators seek a newer, more glamorous panacea.

By describing the R&E visitation program, this book provides development professionals and community leaders with useful information and insights about another tool to use in the never-ending battle of trying to maintain economically healthy communities.

Gary B. Hansen
Utah State University

Moyer, H. Wayne and Timothy E. Josling. *Agricultural Policy Reform: Policy and Process in the EC and the USA*. Ames: Iowa State University Press, 1990, xx + 235 pp., \$27.95.

Moyer and Josling have produced an important text for those interested in policy formation. Their analysis of the problems of decision making in the EC and the United States is insightful and thought provoking. Their book is not so much about "Agriculture Policy Reform" as it is about the "politics and process" that prevent reform or perhaps retard the evolutionary process of policy reform.

In their introduction, they define reform as "a significant shift in policy direction, usually involving changes in instruments, arising from general dissatisfaction with the current operation of the policy" (p. xviii). They admit to the subjectivity of this definition and qualify it by reiterating that reform implies a significant shift in policy.

In part 1, Moyer and Josling review a variety of public choice and operational policy models as a preliminary exercise to the development and presentation of their "Process Schematic Approach" (p. 13) which they base on earlier work by Petit. The models reviewed provide some insight into the decision process, but each is inadequate as a model of agricultural policy reform because of the underlying assumptions

of the model. Rational choice models rely on a calculating central national actor. Public choice analysis assumes the maximization of egotistic utility subject to constraints. In the organizational process paradigm, the best predictor of what an organization will do today is what it did yesterday. The partisan mutual adjustment model assumes that policy is created through bargaining and consensus building.

Graduate students doing required reading should read part 1 carefully, otherwise you may safely skip to either part 2 "Agricultural Reform in the European Community," or part 3, "Agricultural Reform in the United States," to begin reading. If you desire, these sections may be read in reverse order because they are independent analyses of the policy processes in the EC and the United States. An occasional referral to the schematic in table 1.1 or a quick skim of part 1 is all that is required to follow either analysis.

Moyer and Josling hypothesize that: (a) initiative for reform must come from high-level central decision makers, (b) legislative actors would almost never initiate reform, (c) bureaucratic actors are not likely to be in the vanguard of reform, (d) a budget crisis provides the most potent stimulus to change, (e) the normal agricultural policy process will probably not make the necessary reforms, and (f) the process will change by broadening the range of participants or by shifting the decision locus. These hypotheses are not so much tested in the analysis as they are supported by anecdotal evidence in the chronology of policy development for the 1980s. In the end, it is difficult to determine the distinction between hypotheses, pre-conceptions, and conclusions.

Their schematic makes explicit a procedural outline, but it does not really help one identify the absolute or relative importance of the actors and procedures in the policy process. How one weighs the input from various actors remains unclear, and the importance of individual personalities is not really discussed in the Moyer-Josling analysis. Pitting a high-level somewhat naive reform-oriented ideologue against an entrenched group of rent seekers will not likely produce a compromise that will result in change. Yet, rent seekers will accommodate to change if the outcome remains ambiguous enough to convince a majority that they have nothing to lose and possibly something to gain. Or change might occur if rent seekers perceive that they will be worse off if they do nothing.

One is tempted to argue with some of the details of the Moyer-Josling analysis, especially when they say that farm bills are first formulated in the executive branch and are delicately balanced. However, they refute their own statement by pointing out that both farm bills presented by the Reagan administration were dead on arrival when they reached the Hill. In the final analysis neither the 1981 or 1985 acts bore any relationship to the administration proposals. They do not comment on the 1990 act because the book was published prior to its enactment. Perhaps the most significant lapse in the book is the authors' failure to recognize the significance of the policy change in the

1985 act which bases nonrecourse loan rates on a percentage of past prices. Income support is thus shifted from price enhancement to direct payments. And this shifts the burden from consumers to taxpayer and allows a relatively easy assessment of where the transfers go.

Moyer and Josling rely heavily on work by Petit and Bonnen to support their analysis, and one would be well advised to read these references for insight into the Moyer-Josling hypotheses. However, one need not accept either the Bonnen or Petit arguments to appreciate the outstanding chronological summary and analysis of the policy decision process in the European Community and in the United States in the decade of the 1980s.

The book leaves one somewhat unclear about how a general dissatisfaction with the operation of policy is discovered. And, it is ambiguous about criteria for measuring reform. What is it that is perceived as a policy failure that needs reform? Must reform constitute a move away from government action or is it a somewhat more efficient application of a program?

They conclude that "without a crisis it appears that comprehensive change is impossible" (p. 219). They suggest that reform occurs only in response to crisis which may be an excessive accumulation of stocks, larger and highly visible transfers to producers, external shocks including changes in exchange rates, shifts in export demand or, most important, a budget crisis. Although somewhat pessimistic in their conclusions, Moyer and Josling may, in light of the recent GATT round, be close to the mark regarding the opportunities and options for changing agricultural policies in the United States and the European Community.

Robert D. Reinsel
Economic Research Service,
U.S. Department of Agriculture

References

- Bonnen, James T. *Institutions, Instruments and Driving Forces Behind U.S. National Agricultural Policies*. Washington DC: Resources for the Future. The binational symposium on US-Canadian Agricultural Trade Challenges, National Center for Food and Agricultural Policy, 22-24 July 1987.
- Petit, Michel, *Determinants of Agricultural Policies in the United States and the European Community*. Washington DC: International Food Policy Research Institute Res. Rep. 51, 1985.

Pasour, E. C., Jr. *Agriculture and the State: Market Processes and Bureaucracy*. The Independent Institute. New York: Holmes and Meier, 1990, 258 pp., \$39.95.

Agriculture and the State is a readable and comprehensive discussion of U.S. government programs that relate to agriculture. Pasour defines these programs

broadly to encompass not only commodity and food programs but also international trade, crop insurance, credit, conservation, taxation, research, and extension. Each chapter outlines the program and usually includes a simple economic graph understandable to intermediate economics students. The book could be successfully adopted for use in an undergraduate policy class or to supplement a graduate policy class. While the book reads like a text, including summaries at the end of each chapter, general readers can manage the book. Those readers who lack the skill to read the graphics can easily skim these sections without much loss in content; they will find the explorations of policies uniformly understandable throughout the book.

Pasour presents the entire book within a consistent set of arguments: agriculture programs have failed to achieve their purposes, government failure is at least as prevalent as market failure, and, almost without exclusion, private action is preferable to public action. As Pasour states in his preface in the book:

There are two quite different explanations of U.S. farm programs: the "public interest" and income redistribution. The "public interest" approach, which holds that U.S. farm programs benefit the public at large, represents the conventional wisdom in agricultural policy texts. In this book, in contrast, it is assumed that current programs are better explained by income redistribution. That is, it is assumed here that the persistence of and increased expenditures on farm programs can be traced to the success of agricultural interests in using governmental power to transfer income from taxpayers and consumers to farmers, owners of land and other farm assets, government employees, and agribusiness firms that benefit from current farm policies.—(pages xix and xx)

Pasour sustains this argument even with the more difficult cases of food distribution, conservation, public research, and extension. With respect to food programs, Pasour emphasizes implementation difficul-

ties: "There are formidable information and incentive problems inherent in food-stamp, school-lunch and other programs having means tests" (p. 157). With respect to conservation, he emphasizes the appropriateness of private solutions: "When transactions costs, incentive costs, and information costs are taken into account, there is little if any basis for thinking that an increased role of government in protecting land and soil resources is warranted" (p. 212). With respect to public agriculture research, he advocates private funding while criticizing as faulty those studies which conclude that public funding yields high rates of return: "When these 'government failure' problems are taken into account, it will be found that governments usually can best encourage research investment by defining and enforcing property rights so that privately funded research can be profitably undertaken" (p. 225). With respect to extension, Pasour argues that educational activities tend to be mainly private goods: "The public-goods rationale is less and less valid as one moves from basic science to the development and extension of new technology" (p. 219).

The basis for all of these conclusions lies in Pasour's demand that to merely find the market outcome wanting in terms of some objective is not enough; one must show that the government program—as it actually operates in real-world conditions—is better. His arguments are skillful and particularly challenging to strong believers in the wisdom of public action. These believers could sharpen their own analysis by a careful reading of (and sparring with) the arguments of Pasour. However, Pasour does not address opposing arguments in depth; the novice reader will not be exposed to alternative views of the role of government nor to more sympathetic assessments of existing government programs. *Agriculture and The State* is an unapologetic and well-written argument to limit the role of government in all aspects of agriculture.

Sandra S. Batie

Virginia Polytechnic Institute and State University

Ph.D. Recipients by Subject, 1990

Agricultural Economics General; Curricula and Teaching; Extension, Profession

Harvey Brooks, Iowa State University, "Three Studies in Economics."

Jody W. Lipford, Clemson University, "Religious Organization as Non-Profit Firms: A Study in Organizational Theory."

Ronald Young, University of Kentucky, "The Economics of the Family's Long-Term Care Decision."

Agricultural Inputs; Capital, Agricultural Finance; Land Appraisal and Prices; Labor; Human Capital

Kassahun Abede, University of Minnesota, "Factors Affecting the Demand for Farm Inputs in U.S. Agriculture."

Nelson A. Aguilera, Ohio State University, "Credit Rationing and Loan Default in Formal Rural Credit Markets."

Jong Ha Bae, University of Minnesota, "Estimating Dynamic Input Demand Primal and Dual."

Elsworth D. Beach, North Carolina State University, "A Hedonic Analysis of Herbicides: Does Safety Matter?"

Penelope L. Diebel, Virginia Polytechnic Institute and State University, "An Economic Analysis of Low-Input Agriculture as a Groundwater Protection Strategy."

Hermanto, University of Minnesota, "Demand for Inputs and Supply of Rice Under Risk and Selectivity Bias: A Study of Indonesian Farmers."

Fraidoon Hovaizi, University of Massachusetts, "An Analysis of Structural Change in Residential Electricity Demand."

Janice E. Krakar, Iowa State University, "Canadian Agriculture Factor Retention Under Different Policy Regimes."

Arnold W. Oltmans, University of Illinois, "Aggregate Loan Quality Assessment in the Farm Credit System."

Cheol Soo Park, Iowa State University, "General Equilibrium Asset Pricing Model: An Application to Land Markets."

John Manuel Perez-Garcia, Yale University, "An Analysis of Factors Promoting Land-Use Changes in Brazilian Amazonia."

Janet E. Perry, Oklahoma State University, "Returns to Labor from Farm and On-Farm Employment."

Champak P. Pokharel, University of Minnesota, "Portfolio Adjustment and Asset Fixity Under Uncertainty of Agricultural Banks."

Jayachandran Variyam, University of Georgia, "Credit Market Conditions and Farm Financial Performance."

Feng Xu, Washington State University, "An Econometric Study of Contributions of Parcel Characteristics to Agricultural Land Values in Washington: A Hedonic Approach."

Agricultural Products; Demand, Supply, Prices; Food, Consumer, and Household Economics

Matar Zbar Anazan, Mississippi State University, "An Analysis of Least-Cost Fluid Milk Movement Patterns in Alabama, Louisiana, Mississippi, and Tennessee."

Gary W. Brester, North Carolina State University, "Effects of Measurement Error in Disappearance Data on Estimated Demand Elasticities for Meat."

Dabai Chen, Iowa State University, "Plan and Market'(s): A Theoretical Model of the Chinese Grain Economy."

Peter Y. Chen, University of Alberta, "An Analysis of Market Demand for Meats in Canada."

Joshua A. Greenberg, Washington State University, "An Econometric Analysis of the Alaskan King Crab Industry: A Bioeconomic Perspective."

Ramon G. Guajardo-Quiroga, Texas A&M University, "Economic Impact of the Maquiladora Industry in Mexico."

Mark L. Herrmann, Washington State University, "An Economic Analysis of World Salmon Markets: Effects of Salmon Farming."

Danilo C. Israel, Clemson University, "Total, At-Home and Away-From-Home Catfish Consumption in the U.S.: A Dichotomous and Ordered Logit-Probit Analysis."

Pattana A. Jierwiryanpant, University of Florida, "A Study of the European Community Import Demand for Fresh Fruits and Fresh Grapefruit: A Demand System Approach."

Kwang-Yim Kim, Texas A&M University, "General Equilibrium Analysis of Market and Nonmarket Goods."

Jin Soo Kim, Washington State University, "An Econometric Analysis of Consumer's Valuation of Meat Product Characteristics."

Seok-Joong Kim, Purdue University, "An Assessment of Regional Competitiveness in the North American Egg Markets."

Heung-Dong Lee, Oregon State University, "Input Demand and Output Supply Responses When the Production Function Is Heteroskedastic."

Hwang-Jaw Lee, Ohio State University, "Nonparametric and Parametric Analyses of Food Demand in the United States."

Luis F. Macagno, University of Minnesota, "The Nature and Distribution of Gains from Quality Improving Research in a Multimarket Framework: The Case of Barley."

Raja Masbar, University of Kentucky, "Demand for Soybean Stocks Under Uncertainty in Expectations: A Disequilibrium Market Analysis."

Ahmad Muslim, Mississippi State University, "Indonesian Natural Rubber Supply and Demand Analysis and Policy Implications."

M. Muharminto, Oklahoma State University, "Supply Response for Rice in Indonesia."

Clever Mumbengegwi, Washington State University, "An Econometric Analysis of the Maize Sector in Zimbabwe."

Michael M. Ryan, University of Alberta, "An Economic Analysis of the Alberta Sheep Industry."

Mark W. Stephenson, Cornell University, "The Cost of Processing Butter and Nonfat Dry Milk and the Price of Balancing Seasonally Disparate Supply and Demand for Dairy Products in the Northeast."

Syamsurijal, University of Kentucky, "Household Demand for Dairy Products: An Intertemporal and Demographic Variation Analysis."

Joyce Jong-Wen Wann, University of California, Davis, "Imperfect Competition in Multiproduct Industries with Application to California Pear Processing."

Cathy R. Wessells, University of California, Davis, "An Economic Analysis of the Japanese Salmon Market: Consumption Patterns, the Role of Inventories and Trade Implications."

Economic Growth and Development; Planning Models, Fluctuations; Technological Change; Aggregate Production Capacity; Regional and Community; Migration

Ahmad Baijou, Oklahoma State University, "A Stochastic Agricultural Price Analysis Model of the Moroccan Agricultural Sector Structure and Policy Applications."

Rathin Basu, Virginia Polytechnic Institute and State University, "An Analysis of the Relationship between Sectoral Activity, Diversification, and Structural Change in the Economy."

Sei-Kyun Choi, Purdue University, "Agricultural Policy and Economic Growth: The Case of the Republic of Korea."

Kristy D. Cook, Cornell University, "Determinants of Female Time Allocation in Agricultural Households in Southwestern Kenya."

Bruno H. de Frahan, Michigan State University, "The Effects of Interactions between Technology, Institutions and Policy on the Potential Returns to Farming Systems Research in Semi-Arid Northeastern Mali."

Chanyalew Demese, Kansas State University, "In-

dustry Structure in Rural America: The Effect of Change in Industry Structure on Job Loss in the Rural Counties of the North Central Region."

Abdul-Aziz M. Duwais, Oklahoma State University, "The Saudi Agricultural Sector Model: Structure and Policy Applications."

Erwidodo, Michigan State University, "A Study of Efficiency, Farm-Level Input Demand, and Output Supply of Rice Farms in West Java, Indonesia: Analysis of Panel Data."

Shenggen Fan, University of Minnesota, "Regional Productivity Growth in China's Agriculture."

Masahiko Gemma, University of Minnesota, "Reforming Polish Agriculture: Productivity Growth and Market Behavior of Socialized and Private Firms."

Stephan J. Goetz, Michigan State University, "Market Reform, Food Security and the Cash Crop-Food Crop Debate in Southeastern Senegal."

Mohamed K. Habash, Purdue University, "Potential Returns and Constraints to the Adoption of New Technologies in the Mechanized Rainfed Region (Eastern Vertisols) of Sudan."

Louise Haly, Purdue University, "Economic Evaluation of Technological Innovations for Cassava Production in the Farming System of the Ivory Coast."

Jen Hans, University of Kentucky, "Analysis of Impacts of Biotechnology from the Perspective of Contemporary Production Theory."

Patrick A. Jomini, Purdue University, "The Economic Viability of Phosphorus Fertilization in Southwestern Niger: A Dynamic Approach Incorporating Agronomic Principles."

John M. Kerr, Stanford University, "Economic and Institutional Determinants of Agricultural Mechanization in Egypt."

Sang-Mu Lee, Michigan State University, "Implications of Korean Rice, Beef and Feed Grain Policy Under the Transition from a Developing to a Developed Economy—1967 to 1986."

Julie P. Leones, Cornell University, "Sources of Variance in Household Production and Income in a Philippine Upland Village."

Chi Yuan Lin, Ohio State University, "Primary Agricultural Product Demand and Economic Development."

Yir-Hueih Luh, Pennsylvania State University, "Dynamic Linkages, Learning and Productivity Growth."

Mark R. Lundell, University of California, Berkeley, "Marketing and Supply Response in Lithuanian Agriculture."

Robin Marsh, Stanford University, "Diffusion Under Uncertainty: A Mexican Case Study."

Martin I. Meltzer, Cornell University, "Livestock Biotechnology: The Economic and Ecological Impact of Alternatives for Controlling Ticks and Tick-Borne Diseases in Africa."

Abdelmoneim H. Nagheeb, University of Wisconsin, "Extensification of Agriculture and Deforestation in Sudan: An Economic Analysis Under Uncertainty."

Krishna B. Napit, Virginia Polytechnic Institute and State University, "The Economic Potential of Establishing a Poultry Litter Handling Industry."

Mario Niklitschek, University of Maryland, "Economic Growth in Sub-Saharan Africa: The Implication of the Environmental Degradation."

Shihua Pan, Iowa State University, "The Micro-foundations of Mixed System of Planning and Markets: Some Theoretical Considerations and an Empirical Analysis of the Chinese Agriculture."

Jorge Ramirez, University of Minnesota, "The Role of Irrigation Development in the Indian Foodgrain Sector."

Ganesh P. Rauniyar, Pennsylvania State University, "An Econometric Model of Rate of Adoption of Agricultural Technology for Developing Countries."

Alfredo Riesco, Iowa State University, "Adoption of Improved Livestock Production Practices in the Amazon: An Econometric Analysis."

Jaysingh Say, Oklahoma State University, "Analysis of Income Distribution by Caste and Farm Size for a Panchayat (Village System) in the Tarai Region of Nepal by Means of a Social Accounting Matrix."

Farhed Ali Shah, University of California, Berkeley, "Technological Change, Economic Growth, and Exhaustible Resources."

Barry I. Shapiro, Purdue University, "Potential Constraints, Policy and New Technologies in the Niamey Region of Niger."

Rita Sharma, Cornell University, "The New Economics of the Green Revolution: A Study of Income and Employment Diffusion in Rural Uttar Pradesh, India."

Peter L. Stenberg, University of Minnesota, "A University's Relationship to High-Tech Industry and the Regional Economy: The Case of Minnesota and the U.S."

Gregory Traxler, Iowa State University, "The Economy of Crop Management Research in a Post Green Revolution Setting."

Corinne B. Valdivia, University of Missouri, "Impact of Government Policies on the Small Ruminant Sector of Peru."

David Vandembroucke, Iowa State University, "Public Service Applications of the Central Place Structure of Western Guatemala."

Xianbin Yao, Michigan State University, "Market and Farm Household Level Impacts of Grain Marketing Reforms in China: A Case Study in Xinxiang."

Tielu Yu, Mississippi State University, "An Input-Output Analysis of the Mississippi Economy with Emphasis on the Impact of the Post Service Sector."

Orlando M. Da Silva, North Carolina State University, "The International Market for Frozen Concentrated Orange Juice—Prospects for Brazil."

Velton S. Eddings, Washington State University, "An Analysis of International Beef Trade Flows."

Andre Fargeix, University of California, Berkeley, "Growth and Poverty in Stabilization Programs: A General Equilibrium Model with Financial Markets for Ecuador."

Han-Soo Han, Iowa State University, "The Economic Effects of Taxes, Import Tariffs, and the Foreign Exchange Rate on the Domestic Resource Allocation in the Input-Output Relation."

Huei-Ling (Michelle) Hou, University of Georgia, "Policy Analysis of International Trade in Broilers: A Dynamic Game Approach."

Jin-Ock Kim, Iowa State University, "A Time Series Analysis of the Real Exchange Rate Movement in Korea."

Ain-Ding Liaw, Ohio State University, "A Study on the Effects of Exchange Rate Changes on Taiwan Agricultural Markets—An Elasticity Approach."

Dee-Yu Pai, Ohio State University, "Two Country Model of Macroeconomic Linkages to Agricultural Commodity Flows. The Japan U.S. Case."

Chongkook Park, Iowa State University, "Strategic Trade Policy: Subsidies, Time Consistency, and Quality."

Jesus De Los Santos Pineda, Virginia Polytechnic Institute and State University, "The Impacts of Trade and Agricultural Policies in the Dominican Republic: A Sector Programming Approach."

Velupillai Premakumar, Iowa State University, "Measuring the Impacts of Credit Restrictions on the Trade Performance of Debtor Nations."

Diznarda Salcedo-Baca, University of Illinois, "Distributional Effects of the Mexican Agricultural Trade Policies: The Tomato Case."

Hossana Solomon, Auburn University, "Effects of Export Promotion on Import Demand for U.S. Cotton in the Pacific Rim."

Rogierus J. van den Brink, University of Wisconsin, "Pastoralism, Property Rights and Credit: Three Essays on Indigenous Economic Institutions in West Africa."

Charles T. Worley, Ohio State University, "Implications of Liberalized United States-Canada Trade on Regional Production and Consumption of Grain and Livestock."

Fumiko Yamazaki, Purdue University, "Trade Liberalization in the World Soybean Complex in the Presence of Scale Economies with Multinational Processing Firms."

International Economics; Trade; Integration; Business; Aid

Theophilus R. Brainerd, University of Rhode Island, "An Economic Analysis of Seafood Trade in West Africa."

Marketing; Agribusiness; Cooperatives; Transportation; Industry Organization; Vertical Coordination

Brian D. Adam, University of Illinois, "Incorporating Options and Forecast Information in Producer Hog

Marketing Strategies: Conceptual and Measurement Issues."

Lillian A. Brion, University of Nebraska, "Economic Analysis of Marketing Grain Sorghum in the Philippines."

Greg S. Condas, Pennsylvania State University, "Pricing Efficiency and Hedging Performance on the Baltic International Freight Futures Exchange."

Stuart D. Frank, Ohio State University, "The Structure and Performance of the United States Food Manufacturing Industries: Measuring and Analyzing Vertical Coordination."

Garth J. Holloway, Purdue University, "Analyses of Competition in the Food Industries."

David D. Johnson, University of Minnesota, "Dynamic Aspects of Grain Merchandising Decisions."

Shwu-Jen Lin, Pennsylvania State University, "The Cost of Marketing Food Products: Structures, Prediction, and Decomposition."

Shi-Miin Liu, University of Illinois, "Market Efficiency and the Microstructure of Grain Futures Markets Implied by Return Series of Various Time Intervals."

Rakia Moalla-Fetini, University of California, Berkeley, "Storage Arbitrage Condition and Overshooting: An Impossibility Theorem."

Douglas D. Parker, University of California, Berkeley, "The Economics of Marketing Agricultural Product Quality."

Kenneth K. Sanders, Utah State University, "A Study of the Law of Markets As Enunciated by Jean-Baptiste Say."

Dina L. Umali, Stanford University, "The Structure and Price Performance of the Philippine Rice Marketing System."

Mei Zhang, Mississippi State University, "Impact of Transportation Service Rates on Shipping Rice from the Southern Region of the United States to International Markets."

National Resources; Energy; Conservation; Environment; Wastes; Land Use and Tenure; Recreation; Water; Fisheries

Nir E. Becker, University of Minnesota, "Dynamic Supply from a Common Property Resource: Water Diversions from the Great Lakes."

James F. Booker, Colorado State University, "Economic Allocation of Colorado River Water: Integrating Quantity, Quality, and Instream Use Values."

John R. Boyce, University of California, Davis, "Information and Uncertainty in a Fishery: The British Columbia Salmon Fishery."

Ana D. Capistrano, University of Florida, "Macroeconomic Influences on Tropical Forest Depletion: A Cross-Country Analysis, 1967-1985."

Michael D. Creel, University of California, Davis, "Econometric Problems in Recreation Demand Analysis."

Larry L. Dale, University of Hawaii, "Extraction

Cost, Scarcity Rent and Institutional Choice: Thru Reflections of Resource Scarcity."

Carl R. Dillon, Texas A&M University, "An Economic Analysis of the Edwards Aquifer Water Management."

James M. Ferguson, Virginia Polytechnic Institute and State University, "Free Riding, Contribution Behavior, and Public Goods: The Case of the Virginia Nongame Wildlife Tax Checkoff."

Charles N. Gomersall, Cornell University, "Estimating Dynamic Open Access Models of a Fishery, with an Application to the Flemish Cap."

Philip B. Halverson, Washington State University, "Economic Impact of Interruptible Water Markets on Columbia Basin Project Irrigated Agriculture."

Cindy Ann Jacobs, University of Illinois, "The Impact of Institutional Variables on Measures of Resource Scarcity."

Ahmed Jallala, University of Guelph, "An Assessment of the Economic Impact of Ozone on the Agricultural Sector in Ontario."

Huei-Yann Jeng, Ohio State University, "A Recreational Model with Endogenous Duration and Costs."

Miguel A. Lopez-Pereira, Purdue University, "Economics of Sorghum and Soil Erosion Control Technologies for Small Hillside Farmers in Southern Honduras."

Moeketsi Majoro, Washington State University, "Response of Pacific Northwest Irrigated Agriculture to Rising Electricity Prices."

Lindie Nelson, Cornell University, "Private Woodlands as Household Assets."

Kimberly S. Rollins, University of Wisconsin, "Agriculture and Wildlife: From Principal-Agent Theory to a Wisconsin Economic Policy."

Eirik Romstad, Oregon State University, "Pollution Control Mechanisms When Abatement Costs Are Private Information."

Michael D. Rosen, University of California, Davis, "Property Rights and Public Choice in Water Districts: An Application to Water Markets."

Michael L. Taylor, Oregon State University, "Farm-Level Responses to Agricultural Effluent Control Strategies: The Case of the Willamette Valley."

Naomi Zeitouni, University of Rhode Island, "Efficient Management of Groundwater: An Evaluation of Spatially Differentiated Policies."

Production Economics and Management; Risk and Uncertainty

Bruce L. Ahrendsen, North Carolina State University, "Optimal Capital Structure for the Proprietary Firm: An Empirical Analysis of North Carolina Dairy Farms."

Hans A. Andersson, University of Minnesota, "Owner-Operated Land, Fixed Rent Leases and Share-Leasing Arrangements: Some Implications for Asset and Debt Restructuring of Minnesota Farms."

Arunava Bhattacharyya, Utah State University, "Production and Inefficiency."

Duncan M. Chembezi, University of Missouri, "Regional Supply Analysis, Program Participation and Government Policy Variables: An Econometric Investment of Five U.S. Field Crops."

Kamman Cheung, Stanford University, "Production Efficiency Under Different Economic Regimes: The Case of Chinese Cotton Yarn Industry."

Mario F. Crisostomo, Kansas State University, "Risk and Optimal Crop Rotation Portfolios Including Double Cropping Under the 1990 Farm Commodity Programs: An Application of Crop Growth and Market Simulation Models."

David M. Edwards, Texas A&M University, "Firms' Information Acquisition Decisions in a Microeconomic Context."

Dillon M. Feuz, Colorado State University, "Analysis of Colorado Typical Farms Using the Farm Optimization and Financial Analysis System (FOFAS)."

Kenneth H. Foshee, Mississippi State University, "Estimating the Costs of Producing Container Grown Plants with the Assistance of Computer Accounting Software."

Kenneth A. Foster, University of California, Davis, "A Dynamic Econometric Model of Cattle Inventories and Supply in the United States Beef Cattle Industry."

Doo Bong Han, Texas A&M University, "Investment and Expectations in the U.S. Farm Sector."

David A. Harpman, Colorado State University, "Valuing Instream Flow as a Factor of Production."

Michael R. Langemeier, Purdue University, "The Effects of Liquidity Costs and Capital Rationing on Optimal Farm Planning in an Uncertain Environment."

Nicholas Kalaitzandonakes, University of Florida, "Investment Decisions and Supply Response for Perennial Crops: The Case of Florida Citrus."

Philip Kenkel, University of Kentucky, "An Empirical Examination of Expected Utility Based Approaches to Modeling Behavior Under Uncertainty."

Raymond E. Massey, Oklahoma State University, "A Stochastic Evaluation of Swine Breeding Schemes and Management Strategies."

Peter K. Ngategize, Michigan State University, "Economic Impact of Animal Diseases on Production."

Mark A. Peters, University of Florida, "A Mathematical Programming Model of the U.S. Beef Sector."

James W. Prevatt, Clemson University, "South Carolina Fresh Vegetable Enterprises: A Return-Risk Analysis."

Octavio A. Ramirez, University of Florida, "A Stochastic Optimal Control Formulation of the Risk-Balancing Debt Choice Model: A Basis for Generalized Method of Moments and Maximum Likelihood Estimating of Risk Aversion Coefficients."

Anita Regmi, University of Minnesota, "The Value of Information in Integrated Pest Management of Corn

Rootworm and European Corn Borer in Minnesota."

Iain G. W. Shuker, University of Minnesota, "The Economics of Information Production and Use: An Application to Agribusiness Cooperatives."

Robert C. Thompson, Colorado State University, "Costs and Returns and Economies of Size on Colorado Dairy Farms."

John W. Wade, University of Missouri, "An Analysis of U.S. Feed Grain Yields with Respect to Variability Changes, Skewness, and the Greenhouse Effect."

Man-Sik Yoo, University of Minnesota, "An Application of a Linear-Quadratic Regulatory Optimal Control Problem for a Beef Cattle Operator and Its Implications."

Public Issues and Policy; Agricultural Regulations; Taxation; Inflation, Welfare Programs; Poverty; Regional and Community Development; Education; Health

Akhter Uddin Ahmed, Colorado State University, "Food Policy in Bangladesh: An Analysis of Economic Efficiency and Distributive Justice."

Ahmad Z. Baharumshah, University of Illinois, "The Malaysian Rice Policy: Welfare Analysis of Current and Alternative Programs."

Gnanaraj Chellaraj, Purdue University, "Impact of the Indian Food and Nutrition Policy on Foodgrain Consumption and Imports."

Anthony C. Crooks, University of Minnesota, "Modeling the Dairy Cooperatives Sector for Policy Analysis."

Lahsen Esslimi, Purdue University, "Economic Liberalization and Structural Adjustment: The Moroccan Sugar Policy."

Habib M. Fetini, University of California, Berkeley, "Concepts of Power and the Endogenization of the State in Economic Models."

Devon A. Garvie, University of California, Berkeley, "Essays on Regulation Under Incomplete Information."

Snehylata Gupta, Stanford University, "Rural Credit Policies and the Persistence of Overdues in Indian Agriculture."

Mildred A. Haley, Purdue University, "The Effects of Monetary Policy on U.S. Agricultural Markets: An Integrated Capital Market Approach."

Martin A. Johnson, University of Minnesota, "Agricultural Policies as Nash Equilibria."

Jyoti Khanna, Iowa State University, "Theory and Econometric Analysis of State Government Demand for Public Agricultural Research."

Robert A. Kneuper, Clemson University, "Regulation as Risk-Seeking: Analysis of the Airbag and Superfund Debates."

Gwo-Jiun Leu, University of California, Berkeley, "Multimarket Welfare Analysis of U.S. Sugar Policy."

Gary Lynn McBryde, Washington State University,

"An Economic Model of Local Government Soil Conservation Decision Making."

Geoffrey Miller, Stanford University, "Six Papers on Applied Micro-Economic Policy."

Carl Pechman, Cornell University, "Reliability and Power: The Role of Regulation in Increasing the Competition of Markets for Electric Generation."

Aysen Tanyeri-Abur, Texas A&M University, "An Agricultural Sector Analysis of the United States Sugar Import Policy."

Shelley Jo Thompson, University of Minnesota, "Dynamic Games in Agricultural Policy."

Monte Vandever, Purdue University, "Demand for Crop Insurance and Contract Design: A Case Study for Corn in Indiana."

Erna H. K. van Duren, University of Guelph, "An Economic Analysis of Countervailing Duty Law: Cases Involving Agriculture."

Research Methodology; Modeling; Econometrics; Mathematical Programming; Agricultural Data

Michael Carley, Iowa State University, "Knowledge-Based Support Systems for Statistical Software."

Eric K. Kocher, Oklahoma State University, "Finite Sample Properties of Two Pretest Estimators of Distributed Lag Model Coefficients."

Ph.D. Recipients by Institutions, 1990

University of Alberta

Peter Y. Chen
Michael M. Ryan

Auburn University

Hossana Solomon

University of California, Berkeley

Andre Fargeix
Habib M. Fetini
Devon A. Garvie
Gwo-Juin Leu
Mark R. Lundell
Rakia Moalla-Fetini
Douglas D. Parker
Farhed Ali Shah

University of California, Davis

John R. Boyce
Michael D. Creel
Kenneth A. Foster
Michael D. Rosen
Joyce Jong-Wen Wann
Cathy R. Wessells

Clemson University

James W. Prevatt
Danilo C. Israel
Jody W. Lipford
Robert A. Kneuper

Colorado State University

James F. Booker
Akhter Uddin Ahmed
Dillon M. Feuz
David A. Harpman
Robert C. Thompson

Cornell University

Kristy D. Cook
Charles N. Gomersall
Julie P. Leones
Martin I. Meltzer

Lindie Nelson
Carl Pechman
Rita Sharma
Mark W. Stephenson

University of Florida

Ana D. Capitstrano
Pattana A. Jierwiriyanpant
Nicholas Kalaitzandonakes
Mark A. Peters
Octavio A. Ramirez

University of Georgia

Huei-Ling (Michelle) Hou
Jayachandran Variyam

University of Guelph

Ahmed Jallala
Erna H. K. van Duren

University of Hawaii

Larry L. Dale

University of Illinois

Brian D. Adam
Ahmad Z. Baharumshah
Cindy Ann Jacobs
Shi-Miin Liu
Arnold W. Oltmans
Diznarda Salcedo-Baca

Iowa State University

Harvey Brooks
Michael Carley
Dabai Chen
Han-Soo Han
Jyoti Khanna
Jin-Ock Kim
Janice E. Krakar
Shihua Pan
Cheol Soo Park
Chongkook Park
Velupillai Premakumar
Alfredo Riesco

Gregory Traxler
David Vandenbroucke

Kansas State University

Mario F. Crisostomo
Chanyalew Demiese

University of Kentucky

Jen Hans
Philip Kenkel
Raja Masbar
Syamsurijal
Ronald Young

University of Maryland

Mario Niklitschek

University of Massachusetts

Fraidoon Hovaizi

Michigan State University

Erwidodo
Stephan J. Goetz
Bruno H. de Frahan
Sang-Mu Lee
Peter K. Ngategize
Xianbin Yao

University of Minnesota

Kassahun Abebe
Hans A. Andersson
Jong Ha Bae
Nir E. Becker
Anthony C. Crooks
Shenggen Fan
Masahiko Gemma
Hermanto
David D. Johnson
Martin A. Johnson
Luis F. Macagno
Champak P. Pokharel
Jorge Ramirez
Anita Regmi
Iain G. W. Shuker
Peter L. Stenberg

Shelley Jo Thompson
Man-Sik Yoo

Mississippi State University

Matar Zbar Anazan
Kenneth H. Foshee
Ahmad Muslim
Tielu Yu
Mei Zhang

University of Missouri

Duncan M. Chembezi
Corinne B. Valdivia
John W. Wade

University of Nebraska

Lillian A. Brion

North Carolina State University

Bruce L. Ahrendsen
Elsworth D. Beach
Gary W. Brester
Orlando M. Da Silva

Ohio State University

Nelson A. Aguilera
Stuart D. Frank
Huei-Yann Jeng
Hwang-Jaw Lee
Ain-Ding Liaw
Chi Yuan Lin
Dee-Yu Pai
Charles T. Worley

Oklahoma State University

Ahmad Baijou
Abdul-Aziz M. Duwais
Eric K. Kocher
Raymond E. Massey

M. Muharminto
Janet E. Perry
Jaysingh Sah

Oregon State University

Heung-Dong Lee
Eirik Romstad
Michael L. Taylor

Pennsylvania State University

Greg S. Condas
Shwu-Jen Lin
Yir-Hueih Luh
Ganesh P. Rauniyar

Purdue University

Gnanaraj Chellaraj
Sei-Kyun Choi
Lahsen Esslimi
Mohamed K. Habash
Mildred A. Haley
Louise Haly
Garth J. Holloway
Patrick A. Jomino
Seok-Joong Kim
Michael R. Langemeier
Miguel A. Lopez-Pereira
Barry I. Shapiro
Monte Vandever
Fumiko Yamazaki

University of Rhode Island

Theophilus R. Brainerd
Naomi Zeitouni

Stanford University

Kamman Cheung
Snehylata Gupta
John M. Kerr
Robin Marsh
Geoffrey Miller
Dina L. Umali

Texas A&M University

Aysen Tanyeri-Abur
Carl R. Dillon
David M. Edwards
Ramon G. Guajardo-Quiroga
Doo Bong Han
Kwang-Yim Kim

Utah State University

Arunava Bhattacharyya
Kenneth K. Sanders

Virginia Polytechnic Institute and State University

Rathin Basu
Penelope L. Diebel
James M. Ferguson
Krishna B. Napit
Jesus de Los Santos Pineda

Washington State University

Velton S. Eddings
Joshua A. Greenberg
Philip B. Halverson
Mark L. Herrmann
Jin Soo Kim
Moeketsi Majoro
Gary L. McBryde
Clever Mumbengegwi
Feng Xu

University of Wisconsin

Abdelmoneim H. Nagheeb
Rogerius J. van den Brink
Kimberly S. Rollins

Yale University

John M. Perez-Garcia

Marketed Surplus Under Risk: Do Peasants Agree with Sandmo?

Israel Finkelshtain and James A. Chalfant

Using a newly defined notion of aversion to income risk, the behavior of the marketed-surplus producer under price risk is characterized. Unlike the familiar case first examined by Sandmo, output depends on both ordinal preferences for goods and on risk attitudes. Conditions are found that yield an output level under risk that is smaller than under certainty. If these conditions do not hold, both risk and risk aversion may have a positive effect on output. Implications for econometric studies of risk attitudes are considered and illustrated with an example. Finally, we examine the effect of uncertainty on the peasant's long-run equilibrium.

Key words: marketed surplus, peasant households, production under multivariate risk.

Agricultural economists have devoted considerable effort to studying production decisions by farmers in developing countries. Essential to this effort is a thorough understanding of the role of risk in peasants' production decisions. Important insights into this role of risk are provided by the literature on firms' behavior under risk. Perhaps the most influential paper in this area is by Sandmo, who established that a risk-averse firm facing output-price risk produces less output than a risk-neutral firm. Sandmo and the numerous studies that followed made use of the expected utility hypothesis, combined with an assumption that utility depends on only one random variable—the level of final wealth. Other risks affecting utility must be presumed absent or to have no effect on the behavior under study.

An important characteristic of the agricultural peasant household is the consumption of a portion of its own farm product (e.g., Haessel; Renkow; Toquero, Duff, Anden-Lacsina, and Hayami). Unlike the familiar case studied by Sandmo, when subject to output price risk, such a "marketed-surplus" household faces multivariate risk because the price of one of its con-

sumption goods is also random.¹ Thus, the profits risk affecting wealth cannot be modeled in isolation from other risks. As a consequence, a risk-averse household may no longer choose to produce below the profit-maximizing level of output when the price is random.

The literature (e.g., Roe and Graham-Tomasi) has emphasized that maximization of utility defined on wealth alone is not adequate to describe choices under risk in agricultural households. Anderson, Dillon, and Hardaker (p. 76) asserted that "money is not everything; and the consequences of many decision problems are . . . not well represented in terms of only a single attribute such as monetary gain or loss." Wolf (p. 2) observed that "peasants run a household, not a business concern." Similarly, Ellis (p. 102) proposed that "the interaction of consumption and production within the household causes a unique form of decision making which sets peasants apart from any other kind of production unit. . . ." Restrictions can be found that imply separability between production and consumption decisions. However, Fabella, Pope, and Finkelshtain offer examples showing that strong restrictions must be imposed on preferences in two-period models or in cases of multiattribute objective functions. Therefore, one must take explicit account of the process by which

The authors are, respectively, a lecturer, Faculty of Agriculture, Hebrew University of Jerusalem, Rehovot; and an associate professor, Department of Agricultural and Resource Economics, University of California, Berkeley.

Giannini Foundation Paper No. 964.

Financial support from the Giannini Foundation and the Kinamon Foundation, Faculty of Agriculture, Hebrew University of Jerusalem is gratefully acknowledged.

The authors wish to thank Jock Anderson and Offer Kella for valuable suggestions and comments. Any remaining errors are the responsibility of the authors.

¹ Another example of multivariate risk was examined by Herath, Hardaker, and Anderson, in the context of rice production in Sri Lanka in the presence of yield risk. They considered the varietal choice problem when utility depends on both income and subsistence rice consumption.

profits produce utility, through household consumption decisions, in modeling peasants' production decisions.²

These aspects of household behavior are examined below by combining Sandmo's model with the marketed-surplus literature and by utilizing an alternative to the univariate notion of aversion to risk. Following that, the implications for empirical studies of risk attitudes are explored and the long-run entry/exit decision of the household is described. Finally, we examine the potential for expanding the results to include additional sources of risk.

Peasant Behavior Under Price Risk

Virtually all studies of behavior under risk assume that the producer maximizes the expected value of a utility function defined only over income or final wealth. Aversion to risk in this single argument is measured by the curvature of the utility function, the Arrow-Pratt measure of risk-aversion. To model households facing risks in other arguments of the utility function requires a more general objective function and an alternative definition of aversion to risk. These tasks are taken up in this section.

The Household's Objective Function

We assume that the household is engaged in the production of a food crop (x). Consumption decisions concern leisure (l), a portion of the farm output (m), and an aggregate market good (z), representing the consumption of all other goods. Output and leisure are chosen prior to the realization of prices, while consumption goods are chosen in the harvest period when prices are known. The prices of m and z are denoted by p and q and the wage rate is w .

We assume that the household makes consumption and production choices to maximize the expected value of a utility function U defined over z , m , and l :

$$\max_{m,z,x,l} E[U(z, m, l)],$$

² The fact that such households have limited access to insurance or capital markets makes it especially important to examine the effects of the entire set of risks they face. Of course, the same comments apply to any agricultural household, but the problem seems most acute for low-income, marketed-surplus producers. Higher income farmers are less likely to be consuming part of their output, and the other risks they face are less likely to be correlated with the price of their output.

subject to

$$pm + qz + wl = y(x),$$

where y denotes full income. Because the optimal consumption plans may be revised *ex post*, the *ex ante* decision involves only l and x . Substituting the *ex post* optimal plans for m and z into U leads to the variable indirect utility function $V(y, l, p, q)$. Epstein established its duality to U . Hence, the above problem is equivalent to

$$\max_{x \geq 0, l \geq 0} E[V(y, l, p, q)].$$

In the univariate case where only y is random, V can be reduced to the familiar objective function defined on income or final wealth alone. However, if goods' prices are unknown at the time output is chosen, even under flexibility of consumption choices, their joint distribution is likely to affect output.

Epstein showed that the variable indirect utility function is increasing in y and homogenous of degree zero in y , p , and q . Accordingly, the market good will represent a *numéraire*, so q is normalized to be 1 and other monetary variables are measured relative to q . The real wage rate and the prices of other inputs are assumed to be known when the production and labor supply decisions are made. Thus, we concentrate on risks in the relative price of the food crop and in real income, both of which depend on the random variable p .

Using the notion of full income, total income is the sum of initial wealth, labor income $(T - l)w$, and farm profits given by

$$\pi = p \cdot x - c(x).$$

Here, T is total time endowment of the household and c is the cost function. Any fixed costs are assumed to be contained in c . Output is assumed nonstochastic.³ By the additional assumption that the input prices are known, $c(x)$ is deterministic.

An Alternative Risk Premium

Pratt's risk premium measures the maximum amount that an individual would pay to avoid

³ This assumption is not essential but simplifies the analysis and allows comparison with Sandmo's model. The main results hold under more general forms of stochastic profits, provided, of course, that p remains random.

income risk when nothing else is random. In terms of the current model, it is defined as S_p in

$$EV(l, y, \bar{p}) = V(l, \bar{y} - S_p, \bar{p}),$$

where the \bar{p} subscript indicates that the price of m is fixed at its mean. If the price of the portion of output consumed at home remained random, with only income stabilized, Pratt's risk premium would not necessarily measure the willingness to pay for income insurance. The risk premium appropriate for such a situation must be defined as the maximum amount that the individual would pay to avoid income risk when p remains random.⁴

$$EV(l, y, p) = EV(l, \bar{y} - S, p).$$

The interpretation of the modified risk premium S is analogous to that of the regular Pratt risk premium in the traditional univariate model, which appears as a special case, when the consumption price p is fixed at its mean.

This can be illustrated for small risks by a Taylor approximation of the equation that defines S , a derivation of which is given in an appendix, available upon request. Assume that the producer faces uncertainty about income and the price of one consumption good. For small risks, the risk premium is given by

$$S = -\frac{1}{2} \sigma_{yy} \frac{V_{yy}}{V_y} - \sigma_{yp} \frac{V_{yp}}{V_y},$$

where subscripts on V denote partial derivatives, σ_{yy} is the variance of total income, σ_{yp} is the covariance between total income and p , and the partial derivatives are evaluated at (l, \bar{y}, \bar{p}) . The first term in the above expression is Pratt's risk premium, the maximum amount that the producer is willing to pay to stabilize income when prices are fixed. The second term captures the value (or cost) associated with the stochastic interaction between the consumption price and his income. If this covariation did not affect utility (because $V_{yp} = 0$, implying that V is additively separable) or did not exist (due to independence

of y and p , or fixed p), the risk premium would reduce to the univariate measure.⁵ Otherwise, S differs from the Pratt risk premium in magnitude and possibly in sign.

In the case of a marketed-surplus good, this expression can be simplified somewhat. If the price of output is the only random variable affecting profits and consumption, the above expression reduces to

$$S = -\sigma_{pp} x \left[\frac{1}{2} x \frac{V_{yy}}{V_y} + \frac{V_{yp}}{V_y} \right],$$

because the variance of income, σ_{yy} , is simply the variance of revenues from production, $x^2 \sigma_{pp}$, where σ_{pp} denotes the variance of the output price, and the covariance of income and price, similarly, is $x \sigma_{pp}$. Since the sign S is determined solely by the sign of the term in brackets, this expression suggests that it might be possible to find conditions on utility functions whose owners have a risk premium of a particular sign, regardless of the probability distribution of p . These new conditions are needed since $V_{yy} < 0$, the univariate condition, no longer is sufficient for $S > 0$. This is accomplished in proposition 1, which illustrates the relationships between the sign of S and various parameters of the peasant's preferences, including the marketed surplus (the difference between x and m).

PROPOSITION 1. *Let η be the income elasticity of the household's demand for home-consumption (m) of the farm crop. A necessary condition for a positive risk premium S is*

$$\eta > r \left[1 - \frac{1}{2\rho} \right],$$

where ρ is the ratio of home-consumption to total production (m/x) and r is the Arrow-Pratt measure of relative risk aversion. (A proof of the proposition is available from the authors upon request.)

The expression derived for S using a Taylor-series approximation suggests that, for small risks, this condition is both necessary and sufficient.⁶ Also, if the household consumes less (more) than half of its output, a necessary condition for a positive (negative) risk premium is that the farm-produced good is a normal (infer-

⁴ At first glance, such a risk premium seems to attach a value to an impossible stabilization scheme. The price of the output x could be stabilized independently of the consumption price, however. For instance, the government could operate a consumer subsidy scheme where the price to consumers was uncertain and there was a pre-announced price at which all output would be purchased from producers. A similar outcome could occur if the producer forward contracts all of his output at an unbiased futures price and then decides consumption based on the realization of p . In any event, the concept of the risk premium does not require that such a partial insurance scheme actually be offered.

⁵ Independence is sufficient for that to occur only for small risks. For large risks, independence would not insure equality but would guarantee that S and the univariate risk premium would have the same sign.

⁶ A sufficient condition that holds for large risks as well as small risks is $\eta > r > 0$.

rior) one ($\eta > (<) 0$). If the household consumes exactly half of its output, a necessary condition for a positive risk premium is that the farm-produced good is a normal one ($\eta > 0$). Assuming that $r > 0$, as the proportion of output consumed at home increases, the lower bound on η necessary for $S > 0$ increases as well.

We turn now to the main results of the paper, concerning the level of output for the "marketed surplus" producer. To establish these results, the objective function and risk premium defined above are used.

Optimal Output in the Short-Run

The first-order conditions for the household's maximization problem are

$$(i) \quad E[V_l - V_y w] = 0, \text{ and}$$

$$(ii) \quad E\{V_y[p - c'(x)]\} = 0.$$

Condition (i) requires that the optimal allocation of time equates the expected marginal utility from leisure and the expected marginal utility from the wage payment saved by substituting an additional unit of the owner's labor for hired labor. Condition (ii) states that the expected marginal utility from an additional unit of production vanishes. This condition will be used below to derive qualitative results regarding the level of production. The sufficient conditions for a local maximum are assumed to hold and are given by:

$$a_{11} \equiv E[V_{yy}(w)^2 - 2 \cdot V_{yl}w + V_{ll}] < 0$$

$$a_{22} \equiv E\{V_{yy}[p - c'(x)]^2 - c''(x)V_y\} < 0$$

$$a_{11} \cdot a_{22} - (a_{12})^2 > 0,$$

where

$$a_{12} \equiv E\{V_{yy}[p - c'(x)]w - V_{yl}[p - c'(x)]\}.$$

The second necessary condition can be rewritten as

$$E[V_y p] = E[V_y c'(y)], \text{ or}$$

$$E[V_y(p - \bar{p})] = E\{V_y[c'(x) - \bar{p}]\}.$$

The left-hand side is the covariance between the marginal utility of income and the output price. In Sandmo's model, when the producer is risk averse ($V_{yy} < 0$), this covariance is clearly negative. This is because V_y is decreasing in income and hence in p . This implies that the expected price of output exceeds marginal cost ($\bar{p} > c'(x)$).

This observation leads to the conclusion that, in the classic case, output for risk-averse producers is strictly less than the expected profit-maximizing level.

However, in the case of a peasant producer, p affects V_y not only through the income argument (as captured by V_{yy}) but also through the cross derivative V_{yp} . As a result, $V_{yy} < 0$ is not sufficient to ensure Sandmo's result without more information about the peasant's tastes, production technology, and the probability distribution of prices. Hence, the level of production may exceed the optimal level under certainty. Below we derive plausible conditions under which Sandmo's result is preserved and discuss the consequences of the failure of these conditions for eliciting risk aversion or characterizing behavior. Proposition 2 establishes the conditions under which the level of output is adversely affected by risk.

PROPOSITION 2. *Let the optimal quantities produced under certainty and uncertainty be x^c and x^u , respectively. The relationship between x^c and x^u is given by*

$$x^u \leq x^c \text{ for all risks } \Leftrightarrow \eta \geq r \left(1 - \frac{x^u}{m} \right).$$

Proof. The proof is based on the equivalence of statements (i)–(iv), which holds when conditions (i)–(iii) are required to hold for all price risks:

$$(i) \quad x^u \leq x^c$$

$$(ii) \quad \bar{p} \leq c'(x)$$

$$(iii) \quad \text{Cov}(V_y, p) \leq 0$$

$$(iv) \quad \frac{dV_y}{dp} \leq 0.$$

Using Roy's identity, condition (iv) becomes

$$\begin{aligned} \frac{dV_y}{dp} \leq 0 &\Leftrightarrow \eta \geq r \left[\frac{s_m - \beta}{s_m} \right] \\ &\Leftrightarrow \eta \geq r \left[\frac{m - x^u}{m} \right] \\ &\Leftrightarrow \eta \geq r \left[1 - \frac{x^u}{m} \right], \end{aligned}$$

where s_m denotes the share of the marketed-surplus good in total consumption expenditures and β denotes the share of the risky income in total wealth. \square

The intuition behind the awkward looking expression is clearest using the following relationship involving the consumption parameters, derived by rearranging the first condition on η .

$$x^u \leq x^c \Leftrightarrow r\beta \geq s_m(r - \eta) = s_m r(1 - \mu), \text{ or}$$

$$x^u \leq x^c \Leftrightarrow \beta \geq s_m(1 - \mu),$$

where $\mu = \eta/r$. Besley made use of a "profit" function Φ for the consumer, to represent the expenditure required at any price vector to maintain a particular marginal utility of income. Among other things, this function satisfies

$$\frac{\partial \log \Phi}{\partial \log p} = s_m(1 - \mu),$$

which gives the percentage change in expenditures required to maintain marginal utility following a 1% change in p .

The left-hand side of the above inequality, β , has a similar interpretation; it is the percentage increase in wealth from a 1% increase in the price of output:

$$\frac{dy}{dp} \cdot \frac{p}{y} = x \cdot \frac{p}{y} = \beta.$$

Thus, uncertainty about p has no effect on the "marketed surplus" firm (i.e., $x^u = x^c$) when the change in wealth resulting from a price change equals the change in consumption expenditures needed to maintain a constant marginal utility of income:

$$\beta = s_m(1 - \mu).$$

Effectively, the producer is made risk neutral with respect to price risk because marginal utility of income is constant with respect to changes in p . When the wealth effect dominates, Sandmo's qualitative result is preserved. His result is reversed when the consumption effect dominates.

Proposition 2 also suggests that a crucial factor in the relationship between output levels under certainty and uncertainty is the degree of self-sufficiency of the farm—in other words, whether or not there is a marketed surplus. Empirical evidence regarding this parameter of peasants' behavior is provided by Ahmed and Bernard (table 1, p. 17). They reported that 40% of the farms in Bangladesh are deficit farms—net buyers of rice.⁷ Only 30% of the rice farms in this

country always have a positive marketed surplus.

Returning to the proposition, a sufficient condition for Sandmo's qualitative result to hold for the latter type of farms is that the good produced be a normal good ($\eta > 0$); the lower bound on η given in the proposition will be negative, ensuring a Sandmo-like response for risk. On the other hand, for the other 40% of the farms that are net buyers of rice, a positive income elasticity becomes a necessary condition for Sandmo's result to hold while a negative one is sufficient for a reversal of that result.

If the producer does not consume any of his farm product ($s_m = 0$), the level of output is separable from preferences for goods and, just as in Sandmo's model, risk aversion is sufficient for uncertainty about the price of output to affect production adversely. Alternatively, this last result is obtained if the variable indirect utility function is additively separable in income and the price p for consumption of m . However, this condition implies rather restrictive preferences because it requires that $\eta = r > 0$.

Examination of propositions 1 and 2 illustrates an important difference between the cases of univariate and multivariate risk. The condition for a positive risk premium, in the Sandmo framework, is identical to the one for a level of output that is less than under certainty. However, in the current framework, the condition that ensures a smaller level of output under uncertainty differs from the condition for a positive risk premium. Thus, a producer may prefer a stable income and, at the same time, produce more than under certainty.

Figure 1 illustrates the above results. In each of the four graphs, the relationship between the optimal level of output (x) and the coefficient of variation of price (ψ) is given for three levels of aversion to risk, as measured by the Arrow-Pratt measure of relative risk-aversion ($r = 0, 1/2, 1$). The cost function is $1/2x^2$ and the mean price is assumed, for simplicity, to be 1. The four graphs combined show the effects of η , the income elasticity of demand for the marketed-surplus good, and ρ , the ratio of home consumption m to total output x .

In figure 1a, the optimal level of output is shown to decrease with respect to risk for three levels of r . The Sandmo results hold, in the sense that increases in r or in the uncertainty about the output price (as measured by ψ) reduce the optimal output. However, note that output is affected by increasing risk even when the pro-

⁷ The required purchasing power comes, presumably, from other outputs or from off-farm sources of income.

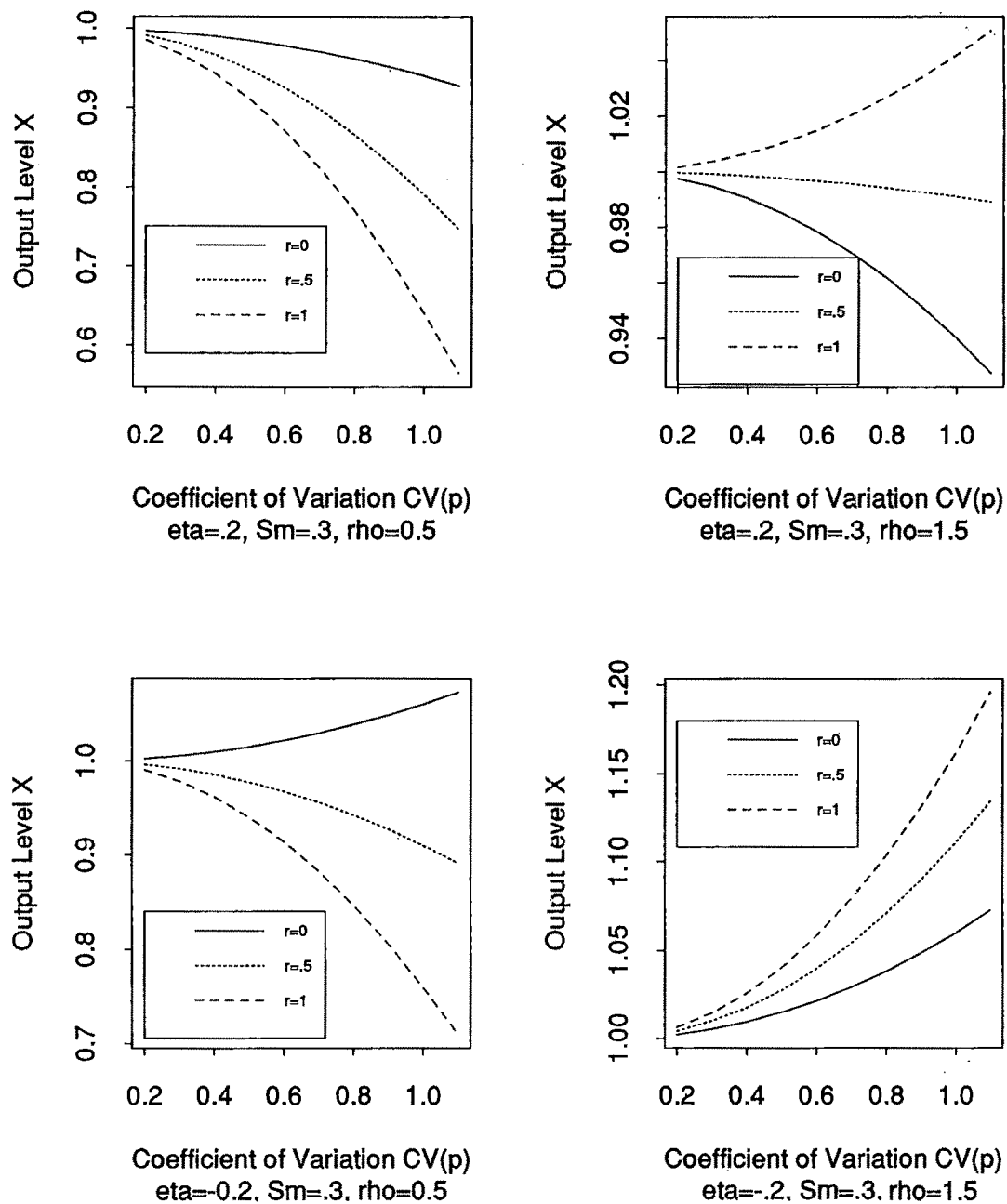


Figure 1. Effects of risk and risk aversion on output levels

lucer is risk neutral. Such a producer is indifferent to the income risk associated with p but not the consumption effect. Returning to the condition

$$\eta > r \cdot \left[1 - \frac{1}{\rho} \right],$$

the risk-neutral producer will produce less than under certainty if $\eta > 0$, more if $\eta < 0$, and will be unaffected only if $\eta = 0$. In figure 1b, everything is the same except the producer is a net buyer of the product; here ρ is 1.5. The risk-neutral and slightly risk-averse producer's behavior is qualitatively the same as

before, while the more risk-averse producer increases output with an increase in risk. Also, note that increases in r are associated with increased output for a given level of risk.

Figure 1c shows that a risk-neutral producer may also increase output as risk increases. In this case, the producers are net sellers, but the good is an inferior good. The Sandmo result concerning increases in r thus holds, but the negative value for η causes the increase in output, as risk increases, to be lessened for risk-averse producers. The risk-neutral or only slightly risk-averse producer actually produces more as risk increases; how risk averse one could be and still exhibit this behavior would vary with the levels of η , s_m , and ρ .

Finally, the combined effects of a negative value for η and being a net buyer are shown in figure 1d. Here, both aspects of the usual responses to risk are reversed. Increases in risk cause all producers to increase output, and the level of output is increasing in r for any level of risk.

Implications for Econometric Studies

Econometric studies of peasants' risk attitudes (e.g., Antle) are usually based on the observed gap between expected revenue and marginal cost. In the traditional model, this gap equals the derivative of the Pratt risk premium with respect to output (e.g., Antle, Flacco and Larson). This last quantity, termed the marginal risk premium, is a function of the observed moments of the probability distribution and the agent's risk attitudes.⁸ This relationship permits the agent's degree of risk aversion to be inferred by a comparison of the above gap to the variance of profits.

However, this relationship is misspecified if the agent's objective function is defined over other random variables as well. The wrong risk premium will have been used. Instead, the derivative of the risk premium S defined earlier is appropriate. For a correct elicitation of aversion to risk in such cases, one needs information about the moments of the joint distribution of the different risks in the objective function and about the agent's attitudes toward the other risks. These may depend, as was shown in the preceding section, on the agent's ordinal preferences.

Suppose that the small risk assumption is valid.

A first-order Taylor approximation of the producer's first-order condition yields

$$\frac{\bar{p} - c'(x)}{\bar{p}} = \psi^2 [r\beta + s_m(\eta - r)],$$

where ψ is the coefficient of variation of the output price. Data on \bar{p} , ψ , production inputs, and output would facilitate the estimation of the coefficient of ψ in the above equation. Previous studies have interpreted the coefficient on ψ^2 or on variance as the Arrow-Pratt measure of risk aversion (e.g., Antle, p. 774). However, in the marketed-surplus case, this coefficient clearly cannot be interpreted as a coefficient of risk aversion because it also includes parameters that describe ordinal preferences for home consumption. Even if it is known that β is around one, so that safe income is negligible compared to the farm income, such an inference is not possible. When $\beta = 1$, one obtains a linear combination of r and η , with the weights determined by the size of s_m ; the larger is s_m , the more the estimate includes the effects of η . And, if the sufficient condition for Sandmo's result holds—that η is larger than r and both are positive—the result is that the estimated degree of risk aversion is too large.

When does this quantity reduce to the Arrow-Pratt measure of risk aversion? Two alternatives exist. First, $s_m = 0$, so the producer does not consume any of his farm product. Second is the case where $\eta = r$, which was discussed in preceding section. As in the traditional model, if these conditions hold, estimation of an individual's attitude toward risk is possible using the observed gap between expected price and marginal cost. However, if $s_m > 0$, so that the peasant consumes a significant portion of the product grown on his farm, and his variable indirect utility function is not additively separable, the traditional estimation procedure yields biased estimates.

In some extremes, the result may even be estimates with the wrong signs, indicating risk-seeking behavior although the agent is, in fact, risk averse. The above expression for the first-order condition reveals that if $s_m > \beta$ and $\eta \leq 0$, the level of output under uncertainty is increasing in the Arrow-Pratt measure of relative risk aversion. Thus, if the producer is a net buyer of the farm product,⁹ and the good is inferior,

⁸ Typically, researchers have assumed small risks, so only the second moment, variance, matters.

⁹ That this might occur seems rather surprising, but the data mentioned earlier from Bangladesh, cited by Ahmed and Bernard, show that it can.

greater risk aversion is associated with a smaller gap between expected price and marginal cost. Hence, in this case if one simply regresses the gap between expected price and marginal cost against variance, the estimated coefficient, mistakenly believed to equal r , will be negative even though the producer is risk averse.

Figure 2 illustrates these findings for some plausible values of consumption parameters. On the vertical axis is the expected value of the estimated coefficient in the simple regression where the dependent variable is the percentage gap between expected price and marginal cost and the explanatory variable is the coefficient of variation of the output price. The horizontal axis gives values of the Arrow-Pratt measure of relative risk aversion.

Ideally, the relationship between these two would be given by a 45-degree line from the origin. The figure illustrates that this is not the case when the estimation procedure is based on a simple univariate Sandmo-type model. Ignoring multivariate risk causes behavior that is partly the result of uncertainty about a consumption price to be ascribed solely to aversion to income risk.

If $\rho = 1$, the Arrow-Pratt coefficient of risk aversion has no effect on the estimate. In fact, the regression coefficient is an unbiased estimate of the product of the income elasticity and the budget share. For deficit farms ($\rho > 1$), one is likely to obtain negative relationships, as the figure shows: increases in the value of r lead to decreases in its estimated value. Only for farms with a positive marketed surplus will greater risk aversion tend to result in larger estimates.

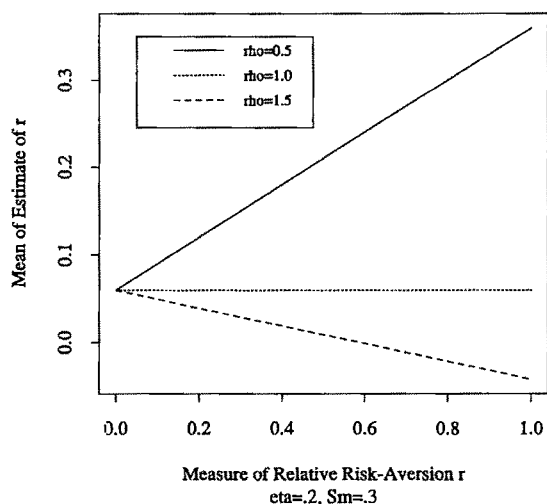


Figure 2. Estimation of risk aversion

Long-Run Entry and Exit Decisions of the Risk-Averse Peasant

The discussion thus far has assumed that the household produces a positive level of output and hence ignores the possibility of quitting farming. However, the effects of price risk may be severe enough to induce farmers to quit the farm and to look for alternative income sources. For example, peasants might choose to migrate from rural areas to urban areas to sell their labor services. This is the topic of the current section, which examines the long-run equilibrium and serves to illustrate how the modified risk premium can lead to different conclusions than the usual risk premium.

Proposition 3 generalizes the results of Sandmo, Paris, and Flacco and Larson, establishing the relationship between the peasant's attitude toward risk and the difference between expected price and average cost which induces entry/exit. Because the cost of production includes the opportunity cost of the peasant's labor endowment (that opportunity cost being the urban wage), the proposition yields the conditions under which it would be optimal to quit farming and migrate to the city.

PROPOSITION 3. *Exit from farming occurs where average cost is smaller (larger) than expected price if and only if the peasant is averse to (seeks) income risk, in the sense that $S > (<) 0$.*

Proof. We prove the result for aversion to income risk; the case of risk seeking can be proved similarly. Let the initial level of wealth be W_0 and profits from farming π . The producer is indifferent between his two alternatives if and only if expected utility is the same whether he produces or not:

$$E[V(l, \pi + W_0, p)] = E[V(l, \bar{\pi} + W_0 - S, p)] \\ = E[V(l, W_0, p)], \text{ or}$$

$$\bar{p} = \frac{C(x)}{x} + \frac{S}{x} = 0. \quad \square$$

Various researchers (cited in Katz and Stark, for instance) have presented surprising evidence that migration has occurred even though "in many cases the expected income in the urban area is not larger than the expected income in the rural area" (Katz and Stark, p. 135). Risk aversion in the Arrow-Pratt sense can explain such migration if income sources in the city are safer. However, urban job markets in developing

countries are often characterized by a high level of uncertainty about employment.

Katz and Stark suggested imperfections in the capital market as a possible explanation. Another alternative is provided by aversion to income risk, as captured by a positive value for our modified risk premium S . Recall that S consists of the Arrow-Pratt premium plus an additional term. This additional term may be positive, providing an additional reason to quit farming. It may turn out, then, that peasants will be willing to bear additional income risks in the urban area—to escape a positive correlation between income and a consumption price—thus reconciling migration with aversion to income risk.

Turning to the industry, in long-run equilibrium each participant in the industry must be indifferent between operating or quitting farming. This requirement implies the following characteristics of the long-run equilibrium, stated as corollary 1, an immediate result of proposition 3.

COROLLARY 1. *If the industry consists of peasants who are averse to income risk, the long-run equilibrium is characterized by positive expected profits.*

The above condition replaces the traditional “zero profits” characterization of the long-run equilibrium of the industry under certainty. Thus, the average risk premium drives a wedge between expected price and average cost. As is true in the case of univariate risk (Flacco and Larson), production under risk, if producers are risk-averse, does not take place where long-run average cost is at a minimum.

Production may also take place even though expected profits are negative. While this could occur in a univariate framework, it would be possible only with risk-preferring producers. Here, it could occur even with risk aversion in the Arrow-Pratt sense ($V_{yy} < 0$), as long as the consumption effect of randomness in p was large enough to yield $S < 0$. To summarize, aversion to income risk, as defined by a positive modified risk premium, is both necessary and sufficient for the peasant to produce only when expected profits are strictly positive, while the Arrow-Pratt measure is neither necessary nor sufficient.¹⁰ An empirical implication of this, once more, is that even if positive expected profits

are observed in the long run, one cannot conclude that producers are risk averse in the sense of Arrow and Pratt. Similarly, negative expected profits (or apparently suboptimal exits from farming) need not imply risk-preferring behavior.

Discussion of Possible Extensions

Uncertainty about prices is indeed one of the important risks with which peasants must contend. For example, Newbery and Stiglitz and Ahmed and Bernard present evidence that in developing countries, the coefficient of variation of the price of a typical agricultural commodity is around 0.5. However, additional risks affecting profits or other arguments of the utility function are typically present.

Uncertainty about the relative prices of other commodities consumed is another risk that may be present in the utility function. It is unlikely that the producer faces output price risk while all other prices in the economy are known or move together (pure inflation), as essentially was assumed in our model. It is straightforward to generalize our results, however. The risk premium S is defined as before, but p now denotes a vector of goods' prices. For small risks, it can be shown that

$$S = -\frac{1}{2} \sigma_{p_x}^2 x^2 \frac{V_{yy}}{V_y} - \sum_{i=1}^N \sigma_{p_x p_i} x \frac{V_{yp_i}}{V_y}.$$

This is a generalization of the risk premium of the second section of the paper, which was defined for the case of uncertainty about the price of output and the price of one consumption good. This expanded risk premium consists of the univariate income-risk premium and a term capturing the individual's aversion to the covariation of income (through the output price p_x) with all other prices p_i ($i = 1, \dots, n$). The latter could be positive or negative.

Both components of S have a derivative with respect to output; as a result, the marginal risk premium, which represents the difference between price and marginal cost, could have a sign opposite to that of the univariate risk premium. However, unlike the special case of the earlier section, where only the price of the marketed surplus good is random, knowing the sign of the marginal risk premium now requires knowledge of the covariances between prices and of preferences (each V_{yp_i} term). Our analysis that considers only the price of the marketed surplus good

¹⁰ The modified risk premium is also useful for interpersonal comparisons. For example, it can be shown that a peasant who is more averse to income risk would require a higher expected price to remain in farming.

is thus a special case of the more general problem in which all prices are random. As long as other goods tend to have small budget shares and the risks in their relative prices have small degrees of correlation with the output price, it might be reasonable to assume that the overall effect on production of these other risks is negligible. Otherwise, at least for small risks, the expression above replaces the risk premium we used. It complicates the interpretation of the conditions for Sandmo's result or for a positive risk premium because information about all goods is required, but it is still S , and not the univariate risk premium, that describes behavior. In fact, the multivariate risk premium S replaces the univariate (Arrow-Pratt) risk premium any time the situation is one of multivariate risks (Karni, Finkelshtain).

Other risks also affect profits. Certainly, peasants are exposed to uncertainty about yields as well as prices. To accommodate this additional uncertainty, the model must be extended in two directions. First, it must be modified to allow the choice variables to be inputs rather than the level of output. Second, the risk premium is more complicated. Assuming small risks and returning to the case that other goods prices are known, the willingness to pay to stabilize income risk, when it consists of both yield risk and output price risk for the marketed surplus good, is given by

$$S = -\frac{1}{2} \sigma_{yy} \frac{V_{yy}}{V_y} - \sigma_{yp} \frac{V_{yp}}{V_y},$$

where σ_{yy} , the variance of revenues, now reflects variation in price and output, as well as their covariance. In addition, the covariance between income and the price of the good, σ_{yp} , is no longer the same as $\sigma_{pp} \cdot x$; hence, this expression cannot be simplified any further. Both the covariance term and the derivative V_{yp} must be signed before the effect of risk can be determined. Since σ_{yp} depends on the level of production, the marginal risk premium is once more affected by the added risk.

This causes some difficulty in establishing results analogous to propositions 1 and 2. These propositions were based on the assumption that only price risk exists. The risk premium S still may be interpreted as before, but because profits now depend on additional risks, the characterization in proposition 1 of preferences that imply $S > 0$ is no longer valid. Moreover, a proposition similar to proposition 2 is hard to derive because, except for extreme cases such

as where profits and the output price are perfectly correlated, it is impossible to find an expression for dV_y/dp_i of condition (iv).

The conditions under which a producer will always behave as a risk averter with respect to output or will always be willing to pay a positive premium to stabilize income are thus much stronger. While more difficult, the problem is not intractable. Either the problem must be attacked for small risks only, with the Taylor approximation of S , or restrictions must be placed on preferences to obtain large-risk characterizations of behavior.

The main conclusions remain valid: the level of output may well be greater than under certainty, and ordinal preferences affect the level of production. In fact, it can be shown that, for any preferences (except the logarithmic utility function), there is always some probability distribution that implies a higher output level under uncertainty. Also, if additional restrictions are imposed on the joint probability distribution of wealth and the price of consumption goods, a broader set of preferences will imply Sandmo's result. For instance, if they are independent (e.g., if the producer receives a certain price for his output but faces yield risk that is independent of the risk in the prices of consumption goods), it is required only that the producer is averse to risk in the Arrow-Pratt sense.

It is important to emphasize that our discussion in the two preceding sections of the paper also remains valid. An empirical analysis that does not explicitly consider the producer's ordinal preferences for goods will yield biased results if interest is in estimating risk attitudes. Also, aversion to a specific income risk, as defined by $S > 0$, would imply positive expected profits in the long run, and a more risk-averse producer will demand a higher expected price to stay in business.

Conclusions

This paper has shown that many of the familiar results from models of output price uncertainty may no longer hold when producers consume a significant share of their own output. In the short run, a peasant producer who is averse to risk may find it optimal to produce either more or less output than the level that maximizes expected profits. The necessary and sufficient conditions for each of these alternatives were derived. The univariate (Arrow-Pratt) measure of aversion to risk is not sufficient to determine

which will occur. However, the alternative risk premium developed in the paper, which measures aversion to risk in income when other random variables affect utility, does provide a characterization of risk aversion and of the level of output.

Implications were drawn for empirical studies that attempt to determine risk attitudes from observed behavior. Estimates or predictions that use univariate models will, in general, be biased for the case of peasants. Finally, in the long run, the average price the peasant producer requires to remain in business may be larger, equal to, or smaller than average cost.

[Received April 1990; final revision received October 1990.]

References

- Ahmed, R., and A. Bernard. *Rice Price Fluctuation and an Approach to Price Stabilization in Bangladesh*. Washington DC: International Food Policy Research Institute, Feb. 1989.
- Anderson, J. R., J. L. Dillon, and J. B. Hardaker. *Agricultural Decision Analysis*. Ames: Iowa State University Press, 1977.
- Antle, J. M. "Nonstructural Risk Attitude Estimation." *Amer. J. Agr. Econ.* 71(1989):774-84.
- Arrow, Kenneth J. "Aspects of the Theory of Risk Bearing." *Essays in the Theory of Risk Bearing*. Amsterdam: North-Holland Publishing Co., 1970.
- Besley, Timothy. "A Definition of Luxury and Necessity for Cardinal Utility Functions." *Econ. J.* 99(1989):844-49.
- Ellis, Frank. *Peasant Economics: Farm Households and Agrarian Development*. Cambridge: Cambridge University Press, 1988.
- Epstein, Larry. "A Disaggregated Analysis of Consumer Choice Under Uncertainty." *Econometrica* 43 (1975):877-92.
- Fabella, Raul V. "Separability and Risk in the Static Household." *S. Econ. J.* 55(1988):954-61.
- Finkelshtain, Israel. "Aversion to Income Risk in the Presence of Multivariate Risks: Theory and Applications." Ph.D. thesis, University of California, Berkeley, 1989.
- Flacco, Paul R., and Douglas M. Larson. "Marginal and Average Risk Premiums in Long-Run Competitive Equilibrium Under Price Uncertainty." Paper presented at AAEA annual meeting, 2-5 Aug. 1987, Michigan State University.
- Haessel, Walter. "The Price Response of Home Consumption and Marketed Surplus of Food Grains." *Amer. J. Agr. Econ.* 57(1975):111-15.
- Herath, H. M. G., J. B. Hardaker, and J. R. Anderson. "Choice of Varieties by Sri Lanka Rice Farmers: Comparing Alternative Decision Models." *Amer. J. Agr. Econ.* 64(1982):87-93.
- Karni, Edi. "On Multivariate Risk Aversion." *Econometrica* 47(1979):1391-1401.
- Katz, Eliakim, and Oded Stark. "Labor Migration and Risk Aversion in Less Developed Countries." *J. Labor Econ.* 4(1986):134-49.
- Newbery, David M. G., and Joseph E. Stiglitz. *The Theory of Commodity Price Stabilization: A Study in the Economics of Risk*. Oxford: Clarendon Press, 1981.
- Paris, Q. "Long-Run Comparative Statics Under Output and Land Price Uncertainty." *Amer. J. Agr. Econ.* 70 (1988):133-41.
- Pope, R. D. "Separability of Consumption and Production Decisions." Dep. Econ., Brigham Young University, 1989.
- Pratt, J. W. "Risk Aversion in the Small and in the Large." *Econometrica* 32(1964):122-36.
- Roe, Terry, and T. Graham-Tomasi. "Yield Risk in a Dynamic Model of the Agricultural Household." *Agricultural Household Models: Extensions, Applications, and Policy*. Baltimore MD: Johns Hopkins University Press, 1986.
- Renkow, M. "Household Inventories and Marketed Surplus in Semisubsistence Agriculture." *Amer. J. Agr. Econ.* 72(1990):664-75.
- Sandmo, Angmar. "On the Theory of the Competitive Firm under Price Uncertainty." *Amer. Econ. Rev.* 61 (1971):65-73.
- Toquero, Z., B. Duff, T. Anden-Lacsina, and Y. Hayami. "Marketed Surplus Functions for a Subsistence Crop: Rice in the Philippines." *Amer. J. Agr. Econ.* 57(1975):705-9.
- Wolf, E. R. *Peasants*. Englewood Cliffs NJ: Prentice-Hall, 1966.

Market Integration, Efficiency of Arbitrage, and Imperfect Competition: Methodology and Application to U.S. Celery

Richard J. Sexton, Catherine L. Kling, and Hoy F. Carman

This paper develops and applies a methodology to test for efficiency of interregional commodity arbitrage. Application of the methodology requires only time-series data on prices for alternative cities, regions, countries, or product forms. Yet, the approach is capable of generating evidence on a number of market parameters including market integration, arbitrage efficiency, the magnitude of marketing margins, product substitutability, and competitiveness of markets. Estimation is based on a switching regression model with three regimes: efficient arbitrage, relative shortage, and relative glut. Results from application of the model to U.S. celery marketing indicated significant departures from efficient arbitrage for both California and Florida celery.

Key words: arbitrage, celery, imperfect competition, law of one price, market integration, price efficiency.

The nature of markets and their role in price determination are central to economics. Geographic markets are particularly relevant to agriculture because agricultural products are typically bulky and/or perishable and areas of production and consumption are separated; hence, transportation is costly. The geographic boundaries of a market are important in the measurement of supply and demand, in price discovery, and in the structure of competition. Despite this importance, Stigler and Sherwin (p. 555) have noted that "the infrequency with which one encounters actual market size determinations outside the antitrust area is surprising and perhaps disquieting." However, a methodology developed recently by Spiller and Huang to define wholesale gasoline markets geographically may contribute importantly to this line of inquiry. This paper extends the Spiller and Huang approach to measure arbitrage efficiency and adapts the methodology to an agricultural markets context.

An application to U.S. fresh celery markets is also provided.

Contemporary studies of economic markets (and definitions of a market) are usually based on the fundamental definitions of Cournot and Marshall. Two regions are in the same economic market for a homogenous good if the prices for that good differ by exactly the interregional transportation cost. Alternative statements of this same arbitrage concept are that (a) the regional markets are integrated (Goodwin and Schroeder), and (b) the law of one price holds between the two regions (Carter and Hamilton; Goodwin, Grennes, and Wohlgenant).

The failure of two or more regions to adhere to the law of one price may be explained by one or more of the following considerations: (a) the regions are not linked by arbitrage, i.e., they represent autarkic markets (Spiller and Huang); (b) there are impediments to efficient arbitrage, such as trade barriers, imperfect information, or risk aversion (Ravallion, Buccola 1983); or (c) there is imperfect competition in one or more of the markets (Stigler and Sherwin, Faminow and Benson). This latter point is relevant because a normative conclusion often drawn from market integration analysis is that observance of the Marshallian arbitrage condition across markets

Richard Sexton is an associate professor, Catherine Kling is an assistant professor, and Hoy Carman is a professor and chair, Department of Agricultural Economics, University of California, Davis.

This is Giannini Foundation Paper No. 962.

Thanks from the authors go to Julian Alston, Richard Green, and two anonymous reviewers for helpful comments and suggestions on an earlier draft of this manuscript.

implies the existence of efficient, competitive arbitrage forces. However, Faminow and Benson have observed recently that this conclusion is justified only for the prototype point-space trading model wherein "effectively, all buyers and sellers [within a region] are located at a single point" (p. 49).

Information on market integration may, therefore, provide specific evidence as to the competitiveness of markets, the effectiveness of arbitrage (Carter and Hamilton), and the efficiency of pricing (Buccola 1983). However, as Faminow and Benson's work demonstrates, it may be difficult to discern the specific cause(s) of observance or failure to observe the law of one price. This point will be developed further in the following discussion.

Previous Approaches

The traditional methodology to study market integration relies on correlations between the prices in pairs of regions (Richardson, Horowitz). For example, a typical regression model to test for short-run market integration might be

$$(1) \quad P_t^i = \beta_0 + \beta_1 P_t^2 + \beta_2 T_t + e_t$$

where P_t^i , $i = 1, 2$ is the price in region i at time t for a homogenous good, T_t is the transactions cost at t required to ship a unit of the good between the two regions, and e_t is a random error term. A test for short-run integration is provided by the joint hypothesis:

$$(2) \quad H_0: \beta_0 = 0, \beta_1 = \beta_2 = 1.0.$$

Because common shifts in supply and demand may induce substantial correlation among the P_t^i even without interregional arbitrage, analysts have considered variations of equation (1) based upon correlations of price differences (Stigler and Sherwin, Carter and Hamilton).

Researchers have also become increasingly concerned with distinctions between short- and long-run market integration. As Ravallion argues: "in many settings it will be implausible that trade adjusts instantaneously to spatial price differences, so one would be reluctant to accept short-run market integration as an equilibrium concept" (p. 103). The model proposed by Ravallion to test for short- versus long-run integration involves correlating price in one region with lagged own prices and contemporaneous and lagged prices in another region. For example,

$$(3) \quad P_t^i = \sum_{i=1}^n \beta_i P_{t-i}^1 + \sum_{j=0}^n \gamma_j P_{t-j}^2$$

The test of short-run integration is the hypothesis $\gamma_0 = 1.0$, $\gamma_j = 0$, $j \neq 0$, and $\beta_i = 0$ for all i , while the corresponding long-run integration condition requires that $\sum_{i=1}^n \beta_i + \sum_{j=0}^n \gamma_j = 1.0$.

The recent contributions of Goodwin and Schroeder and Goodwin, Grennes, and Wohlgenant are in the same spirit. As Goodwin and Schroeder note (p. 174): "Interregional trade takes time to arrange and to complete. Because of delivery lags, trade between regions is conducted based upon *expectations* of future market conditions" (emphasis added). Thus, these authors construct rational expectations models of arbitrage wherein the market integration condition is one of transportation-cost-adjusted equality between current price in the exporting region and expected future price at the delivery time in the importing region.

However, in competitive markets with continuous trade, failure to observe short-run integration of prices implies normative inefficiency in the arbitrage process. Thus, although long-run integration may be regarded as the more realistic equilibrium concept, short-run integration remains a normative benchmark for market evaluation.

Also short-run market integration may be a more realistic equilibrium condition than the preceding authors imply. For example, in the common trade scenario where one region or country is a dominant exporter shipping to many importers, the flow of product from the exporter to the various importers is nearly continuous. Thus, the efficient arbitrage response to a relative price increase for importer A is not necessarily new shipments to A from the exporter but, rather, a rerouting of shipments from other destinations to A. Depending upon the geographic constellation of importing regions/countries, this adjustment process may occur very quickly, making the law of one price a realistic short-run equilibrium condition.

Faminow and Benson's concern is not the methodology of market integration analysis but rather the interpretation of estimation results, which they show hinges critically on the assumed model of spatial competition. Most prior work has been interpreted from the perspective of the standard point-space trading model (Takayama and Judge) common to international or interregional trade analyses. Here, production and consumption regions are geographically

separate and competition is usually considered only in so far as it emanates from multiple sellers (buyers) located at the production (consumption) point.

Benson and Faminow note, however, that the intraregional trading model (e.g., Greenhut, Norman, and Hung) may be more appropriate in arbitrage analysis for some agricultural commodities. This model embodies a spatial distribution of both buyers and sellers along, say, a line or appropriate two-dimensional surface. Competition is then usually from spatially dispersed buyers and/or sellers, although multiple buyers/sellers may be located at a given point.

The key Faminow and Benson conclusion is that estimation results will be interpreted quite differently from a point-space trading model perspective than from the perspective of spatially dispersed buyers and sellers. For example, short-run market integration may imply anti-competitive base-point pricing rather than competitive FOB pricing and efficient Marshallian arbitrage.¹

The main inference from these recent studies is that the arbitrage model must be chosen and interpreted carefully relative to the trade and market structure characteristics of the product under study. Is short-run (contemporaneous) integration a realistic equilibrium condition? Is the point-space trading model an appropriate characterization of the market(s), or is the model with spatially dispersed agents a more accurate description (Faminow and Benson)?

The Spiller and Huang (SH) model is directly relevant to those situations when integration of markets based on contemporaneous prices is a realistic equilibrium condition, although the methodology is extended in this paper to study lagged price effects. In this context the SH approach overcomes a number of lingering methodological problems in arbitrage models: (a) The arbitrary choice of the price in one region as predetermined is avoided. (b) Transaction costs are estimated within the model. (c) Integration is not treated as an "all or nothing" proposition because regions often may be linked by arbitrage in some periods, but not others, depending upon the supply-demand conditions in each region at time t , as well as the magnitude of T_t .

We turn now to describe the basic SH model. With the extensions proposed here, the ap-

proach is capable of addressing a number of market integration questions relevant to agriculture. Our application to fresh celery markets provides an illustration and helps to highlight some of the points raised by Faminow and Benson.

The Basic Model and Extensions

The SH model is constructed in the point-space tradition and applicable to separate regions that have their own supplies and demands for the product in question. The product may be shipped at a cost across regions. Because each region has its own supply and demand, it is possible to identify autarkic prices in each region. SH express the reduced-form equations of autarky prices for two regions at time t as

$$(4) \quad P_t^{1A} = \pi^1 + \epsilon_t^1,$$

$$(5) \quad P_t^{2A} = \pi^2 + \epsilon_t^2,$$

where the π^i are constant means and the ϵ_t^i represent random shocks to the markets. Given free trade across regions, the actual prices, P_t^1, P_t^2 , may differ from the autarky prices. Specifically, if $|P_t^{1A} - P_t^{2A}| < T_t$, no profitable arbitrage opportunities exist and

$$P_t^1 = P_t^{1A}, P_t^2 = P_t^{2A}.$$

However, if the autarky price difference exceeds T_t , SH assume that competitive interregional shipments will take place until the observed prices in each region differ by exactly T_t . In this case the markets are integrated, and

$$P_t^2 - P_t^1 = T_t > 0,$$

where for simplicity region 2 is assumed to be the higher-price region.

Let transactions cost T_t be modeled as a random variable with constant mean, T :

$$(6) \quad T_t = T + v_t, \text{ where } v_t \sim N(0, \sigma_v^2).^2$$

The probability at time t of no arbitrage opportunities between the two regions (i.e., the probability of autarkic markets) is the constant, λ , where

$$\begin{aligned} \lambda &= \text{Prob}\{P_t^{2A} - P_t^{1A} < T + v_t\}, \\ &= \text{Prob}\{(\pi^2 - \pi^1) + (\epsilon_t^2 - \epsilon_t^1) - v_t < T\}. \end{aligned}$$

¹ Base-point pricing is a collusive strategy whereby all sellers offer the same delivered price schedule consisting of price at a given location (the basing point) plus transportation costs from the basing point to other destinations.

² Constant mean transactions costs over time represents a fairly strong assumption of the SH model and implies that empirical applications of the model require relatively short time series and/or periods of relatively stable prices.

It follows that λ is a function of π^1 , π^2 , T , σ_v^2 , and the distribution parameters of the ϵ_i^j . The probability of observing binding arbitrage (market integration) at time t is $1 - \lambda$, where $\lambda \approx 0$ signifies regions that are almost always integrated (in the same Marshallian economic market).

The above model may be expressed as a switching regression system and estimated using maximum likelihood methods. With the SH approach (a) prices are not treated as predetermined variables in the regression model, (b) interregional shipment costs are endogenous and estimated within the model, and (c) the probability that markets are integrated is allowed to vary continuously.

Extensions

Production of many agricultural products is concentrated in only one or a few regions and, hence, does not conform directly to the SH paradigm, wherein regions have indigenous supply sources. For example, U.S. supplies of many fruits and vegetables, depending upon the season, emanate mainly from California and/or Florida. In these cases shipments typically flow from the producing region to various terminal markets across the country, and the concept of regional autarkic markets is of limited relevance.

A question that arises in these agricultural settings is whether product allocation from the producing region to consuming regions takes place efficiently or whether periodic gluts or shortages occur as the result of product misallocations (Buccola 1983, 1985; Berger et al.). Unlike the SH model, episodes of arbitrage failure cannot be attributed to autarkic market equilibria in these cases. Rather, they must be explained by inefficiencies in arbitrage resulting from trade barriers, imperfect information, and risk aversion, or by imperfect competition.

The extension of the SH model developed in this paper is applicable to many agricultural products. These model applications will be characterized by markets logically linked by arbitrage, with tests for the efficiency and/or competitiveness of the arbitrage process. Test results will usually allow the analyst to draw inferences regarding the efficiency of arbitrage with probabilities of gluts and shortages for a given area, the physical dimensions of the market, and product substitutability. For parsimony of presentation, however, the analysis is developed in

the specific context of the application to U.S. celery marketing.

California is the major celery supplier to U.S. markets throughout the year, but the competitive structure varies by season. Shipments data confirm that California supplies some 80% to 85% of the celery to U.S. terminal markets during the summer and fall months (roughly June through November). Nearly 90% of the California crop is shipped via trucks, with most of the remainder shipped by piggyback van. Michigan is the largest of the secondary summer-fall suppliers, with minor supplies also originating in New York, Ohio, Washington, and Canada.

California's share declines, however, during the winter and spring (December through June) when Florida enters as a major producer, accounting for about 25% of total supplies. Although Florida ships celery to all major terminal markets except those on the West Coast, California celery is preferred to the Florida product because of higher quality, more dependable supply, and California shippers' ability in many cases to provide a single source for many fresh vegetables (Berger et al.).

Berger et al. estimate that there are ninety celery growers and about forty large-scale handlers and marketers of celery in California, with none having a large market share. Moreover, there is no centralized mechanism to coordinate shipments in California. This structure contrasts sharply with the Florida scenario, where only about fifteen large grower handlers operate and marketing is regulated by a federal marketing order featuring flow-to-market provisions. Also, nearly all Florida celery growers belong to a centralized marketing organization, the Florida Celery Exchange (Berger et al.). Contracts with growers give the exchange title to the celery and complete control over its marketing (Kilmer).

Market structure analysis indicates strongly that California celery is allocated competitively. Although Faminow and Benson note that spatial markets often tend to be imperfectly competitive and provide incentives for sellers to price discriminate,³ a large number of sellers located at the same production point renders such discrimination infeasible and competitive FOB pricing emerges in equilibrium. (A formal derivation of this result is provided in Greenhut, Norman, and Hung, chap. 8.)

³ Spatial price discrimination occurs when "the difference in delivered prices between any pair of markets is not equal to the difference in transports costs incurred by the firm in supplying those markets" (Greenhut, Norman, and Hung, p. 102).

Because celery marketing is much more coordinated in Florida than California, it is important to ask whether Florida shippers exercise market power and practice price discrimination across consuming markets or whether pervasive competition from California mitigates these opportunities.

Focusing attention first on celery marketing during the summer months, suppose initially that product flows from California are efficient. Then denoting California FOB prices as p_c and wholesale terminal market prices with the corresponding upper case P^i $i = 1, \dots, n$, we have the following equilibrium arbitrage condition for a given time period t :

$$(7) \quad p_{c,t} = P_t^1 - T_t^1 = P_t^2 - T_t^2 \\ = \dots = P_t^n - T_t^n$$

Departures from (7) imply profit opportunities that, given a perfectly competitive supply, should trigger product reallocations from low- to high-price terminals to restore the equality.

However, if product is not allocated efficiently due to risk factors, imperfect information, significant shipment lags, etc. (see generally Buccola 1983), then periodic gluts or shortages may appear in the various terminals and (7) will not hold for all t . For example, Berger et al. speculate that California causes periodic gluts in Eastern markets during summer

$$(8c) \quad P_t^i - p_{c,t} = T^i + v_t^i - u^i \quad \text{with prob } \lambda_2.$$

Here u is a positive random variable so that (8b) defines a regime wherein the wholesale terminal price exceeds the FOB price plus transactions costs—a relative shortage situation in that less product was allocated to the terminal than indicated by the efficient arbitrage condition (8a). Alternatively (8c) corresponds to a market glut, where the terminal price is depressed below the FOB price plus transactions costs because of excess shipments relative to the efficient arbitrage standard.

The equations in (8), thus, define a switching regression model with three regimes: efficient arbitrage, shortage, and glut. To estimate the model, the likelihood function may be formulated as follows:

$$(9) \quad L = \prod_{i=1}^n [\lambda_1 f_t^1 + \lambda_2 f_t^2 + (1 - \lambda_1 - \lambda_2) f_t^3],$$

where f_t^1 , f_t^2 , and f_t^3 are, respectively, the density functions of (8a), (8b), and (8c), and n is the number of observations. To specify these densities, assume that u_t is distributed independently of v_t with a half normal distribution, i.e., an $N(0, \sigma_u^2)$ distribution truncated from below at zero. Then define $Y_t = P_t^i - p_{c,t}$ and express the densities as follows:⁴

$$f_t^1 = \left[\frac{2}{(\sigma_u^2 + \sigma_v^2)^{.5}} \right] \phi \left[\frac{Y_t - T}{(\sigma_u^2 + \sigma_v^2)^{.5}} \right] \left[1 - \Phi \left[\frac{-(Y_t - T)\sigma_u/\sigma_v}{(\sigma_u^2 + \sigma_v^2)^{.5}} \right] \right], \\ f_t^2 = \left[\frac{2}{(\sigma_u^2 + \sigma_v^2)^{.5}} \right] \phi \left[\frac{Y_t - T}{(\sigma_u^2 + \sigma_v^2)^{.5}} \right] \left[1 - \Phi \left[\frac{(Y_t - T)\sigma_u/\sigma_v}{(\sigma_u^2 + \sigma_v^2)^{.5}} \right] \right], \\ f_t^3 = \frac{1}{\sigma_v} \phi \left[\frac{Y_t - T}{\sigma_v} \right],$$

months by making large shipments with “no pre-arranged destination or price” (p. 35).

To extend the SH methodology to test for inefficient product allocation, we define three regimes that exhaust the possible arbitrage conditions between the producing region and any terminal market i :

$$(8a) \quad P_t^i - p_{c,t} = T^i + v_t^i \quad \text{with prob } 1 - \lambda_1 - \lambda_2,$$

$$(8b) \quad P_t^i - p_{c,t} = T^i + v_t^i + u^i \quad \text{with prob } \lambda_1,$$

where $\phi(\)$ denotes the standard normal density function, and $\Phi(\)$ denotes the corresponding cumulative distribution function for the standard normal. The parameters T , λ_1 , λ_2 , σ_v^2 , and σ_u^2 can be estimated by maximizing the log of (9).

During Florida's winter production cycle, the Florida analogue to the California efficient ar-

⁴ The functions f_t^1 and f_t^2 are best understood through their close analogies to frontier functions (Aigner, Lovell, and Schmidt). Specifically, f_t^1 corresponds to a stochastic frontier equation with a positive error (e.g., a cost function), while f_t^2 corresponds to a stochastic frontier equation with a negative error (e.g., a production function).

bitrage condition, (7), can be formulated and also tested via (9), but these results must be interpreted in a manner consistent with the spatial market structure for Florida celery (Faminow and Benson). Spatial price discrimination is a concern given Florida's marketing order, Celery Exchange, and insulation from nearby spatial competitors. Although Taylor and Kilmer concluded that "the weight of empirical evidence . . . suggests that the Florida celery industry has not enhanced price above [the competitive level]" (p. 41), this conclusion was derived from an aggregate conjectural variations model of oligopoly which may fail to capture spatial market power in specific consuming regions.

Thus, to explore this issue further using an arbitrage model framework, note that a price discrimination equilibrium for Florida would be

$$(10) \quad MR_{F,t}^1 - T_{F,t}^1 = MR_{F,t}^2 - T_{F,t}^2 \\ = \dots = MR_{F,t}^n - T_{F,t}^n \leq p_{F,t},$$

where the $MR_{F,t}^i$ denote marginal revenue from Florida celery sales to the various terminal markets i at time t , and $T_{F,t}^i$ are the corresponding per unit transactions costs. The Florida grower price, $p_{F,t}$, represents a blend price based on revenues from the various terminal markets. Marginal revenue in each terminal market is derived, in turn, in the usual fashion from the residual demand facing Florida shippers in that market.

Florida's residual demand is essentially the market celery demand less the competing California supply at each price.⁵ If Florida has no market power, residual demand for its celery is perfectly elastic, $MR_{F,t}^i = P_{F,t}^i$, for all i , and (10) collapses to the competitive equilibrium condition, (7). Two factors suggest, however, that Florida may face relatively inelastic residual demands in those consuming regions nearest to it: (a) Florida has a decided transportation cost advantage over California in these markets and, hence, is less vulnerable to offsetting California shipments, (b) for most functional forms, demand becomes more elastic as a function of distance from the producing to consuming region (Greenhut, Norman, and Hung). In these cases, price discrimination will favor more distance markets in that delivered price to these markets

will not reflect the full incremental transport costs (Greenhut, Norman, and Hung).

To incorporate spatial price discrimination into an SH-type arbitrage model, define M_t^i as an incremental mark-up to the price, P_t^i , at terminal market i in addition to the transactions costs, T_t^i . It follows that $M_t^i > 0$ implies a price mark up over the FOB price and $M_t^i < 0$ implies freight absorption. The equation for delivered Florida price to any terminal market i is thus:

$$(11) \quad P_{F,t}^i = p_{F,t} + T_{F,t}^i + M_t^i.$$

The modified SH model summarized in (8), and (9) cannot formally distinguish the two components of the $P_{F,t}^i - p_{F,t}$ price difference in (11). That is, the coefficient, $T_{F,t}^i$ from maximizing the log of (9) will include both $T_{F,t}^i$ [see eq. (6)] and a sample average for M_t^i . However, if reasonable inferences about the relative magnitudes of the $T_{F,t}^i$ can be made across the various spatial markets, perhaps inferences about spatial price discrimination can be generated from an SH arbitrage model.

Price Premia

A final extension of the SH model with a direct application to celery marketing concerns tests of the product range.⁶ Although agricultural products often are considered homogenous, subtle, but possibly important, product differentiation may exist among nominally similar products. Examples are similar products grown in different regions, different varieties of a particular crop, or meat products with variable characteristics, such as the degree of marbling.

Two products or varieties of a product belong in the same product market if they are close substitutes (Stigler and Sherwin). The most general approach to measurement of substitutability is estimation of cross elasticities of demand, but data deficiencies and econometric problems may render this method impractical (Baker and Bresnahan), especially for products that are close substitutes.⁷ Yet, specific evidence on the substitutability of products believed a priori to be close substitutes will be important when precise product and market definitions are at issue. (For

⁵ This definition of residual demand is technically correct only for the case where Florida and California celery are perfect substitutes. As noted, some evidence, including our own empirical results, suggests that buyers view Florida and California celery as somewhat differentiated products. Baker and Bresnahan derive residual demand functions for the differentiated products case.

⁶ Stigler and Sherwin have also observed that tests of geographic market definition may also be applied directly to questions of product substitution.

⁷ For example, the closer the substitute relationship between two products, the more colinear their prices will be, making it impractical to include both prices as explanatory variables in demand function estimation.

example, see Hayes, Wahl, and Williams for an attempt to test perfect substitutability in an AIDS framework.)

The SH methodology may also be extended to address a hypothesis that two goods are perfect substitutes (in the same product market). The key point is that a stable relationship must exist between the prices of two perfect substitutes. Specifically, we define perfect substitutability as the existence of a stable premium (possibly zero) separating the prices of two products in our application California and Florida celery. That is,

$$(12) \quad P_{C,t}^i - P_{F,t}^i = \alpha^i + \mu_t^i, \quad i = 1, \dots, n,$$

where $\alpha^i \geq 0$ represents the price premium, and $\mu_t^i \sim N(0, (\sigma_{\mu}^i)^2)$. If California and Florida celery are perfect substitutes, then during all periods when both regions ship to a consuming area, arbitrage forces will ensure that the two prices diverge by the premium, α^i .⁸ However, if the two products are not perfect substitutes, then their price difference may systematically range above or below α^i in response to variations in each product's individual supply-demand factors.

The three regimes of our model of geographic arbitrage efficiency correspond exactly to the model to test for stable price premia (perfect substitutability). That is, we can define three regimes to exhaust the arbitrage possibilities across alternative product forms:

$$(13a) \quad P_{C,t}^i - P_{F,t}^i = \alpha^i + \mu_t^i \quad \text{with prob. } 1 - \lambda_1 - \lambda_2$$

$$(13b) \quad P_{C,t}^i - P_{F,t}^i = \alpha^i + \mu_t^i + u^i \quad \text{with prob. } \lambda_1$$

$$(13c) \quad P_{C,t}^i - P_{F,t}^i = \alpha^i + \mu_t^i - u^i \quad \text{with prob. } \lambda_2,$$

where all variables are defined as in (8). The hypothesis $\lambda_1 = \lambda_2 = 0$ tests whether the products are perfect substitutes, and, qualitatively speaking, the magnitude of $\hat{\lambda}_1 + \hat{\lambda}_2$ is inversely related to the strength of the substitute relationship. Conditional upon rejecting $\lambda_1 = \lambda_2 = 0$, differences in $\hat{\lambda}_1$ and $\hat{\lambda}_2$ can be analyzed to determine possible asymmetries in the substitute relationship.

⁸ For example, if $P_{C,t}^i - P_{F,t}^i > \alpha^i$, then buyers will bid up the price of Florida celery relative to California celery to restore the equality in (12).

Estimation

The switching regression models were estimated using weekly price data for U.S. celery for the four-year period, January 1985 through December 1988. Shipping point and terminal market prices were obtained for both California and Florida celery from various USDA Federal-State Market News Service reports. All prices were in terms of dollars per crate.⁹ California's production is year around, but the producing area shifts seasonally. Thus, Central Coast prices were used during the summer and fall, and South Coast prices were used in the winter and spring. The Florida analysis was based on data from Florida's December through May production period.

Two criteria were used in selecting terminal markets for inclusion in the study: (a) significance of the terminal as a receiver of celery, and (b) location of the terminal. The five largest U.S. terminal markets, Boston, Chicago, Los Angeles, New York, and San Francisco, were included under the first criterion, and Atlanta was chosen under the second.¹⁰ However, Florida does not ship celery to the West Coast, so the analysis for Florida shipments was limited to Chicago, Boston, New York, and Atlanta.

California Celery

The maximum likelihood estimation was conducted using the DFP routine in the GQOPT package. Results from estimation of (9) for California-origin celery are summarized in table 1. Most of the estimated coefficients are statistically significant. The probability, $1 - \hat{\lambda}_1 - \hat{\lambda}_2$, of efficient arbitrage is high, not surprisingly, in Los Angeles (0.76) and San Francisco (0.83), those cities nearest the production regions. However, in all cases the $\hat{\lambda}_i$ are either individually or jointly significant causing a rejection of the hypothesis that arbitrage is efficient or that the law of one price holds for all t .¹¹

The less efficient arbitrage in the eastern markets is consistent with prevailing views on price

⁹ The standard celery crate weighs 60 pounds. The number of stalks per crate varies with the size of the celery. Our prices were usually based on 2.5 dozen stalks per crate. Market News prices are often reported in ranges. In these cases we chose the lower bound of the price range for analysis. The rationale for this choice is provided by SH (p. 137).

¹⁰ Industry experts suggested that about 60% of fresh celery is marketed through the terminals, with the rest contracted directly to major retailers.

¹¹ For Atlanta and Boston joint significance of the λ_i was established based on the usual likelihood ratio tests.

Table 1. Parameter Estimates for California Celery Arbitrage

	Los Angeles	San Francisco	Chicago	Boston	New York	Atlanta
T	2.35 (28.28) ^a	3.83 (32.66)	5.45 (1.86)	4.84 (5.82)	4.53 (11.02)	6.31 (23.13)
σ_v^2	0.62 (5.28)	1.28 (5.39)	0.33 (1.68)	0.87 (2.14)	1.05 (3.24)	0.94 (1.74)
σ_u^2	3.94 (3.52)	7.97 (2.29)	2.23 (5.61)	2.65 (2.42)	3.53 (4.46)	3.20 (1.20)
λ_1	0.15 (2.27)	0.12 (1.80)	0.55 (3.69)	0.78 (1.69)	0.61 (2.46)	0.00 (0.00)
λ_2	0.09 (16.12)	0.05 (1.27)	0.26 (2.60)	0.04 (0.63)	0.03 (0.95)	0.18 (1.31)
$1 - \lambda_1 - \lambda_2$	0.76	0.83	0.19	0.28	0.36	0.82
Log likelihood	-323.44	-379.47	-367.42	-373.22	-391.90	-193.15
Observations	209	209	209	209	209	122
T as % of mean price	28	38	46	41	40	52

^a t -statistics are in parentheses.

ing efficiency (Buccola 1983, 1985). For example, the risk in making unconsigned shipments likely increases with shipment distance due to time lags, possible loss in quality, etc. Quantity and quality of information on market conditions may also decline as a function of distance between shipping point and receiving point. The Atlanta result, $1 - \hat{\lambda}_1 - \hat{\lambda}_2 = 0.82$, differs from the results for the other eastern terminal markets but is explainable by Atlanta's close proximity to Florida production. That is, through the substitute relationship, a glut or shortage of California celery can be corrected by an adjustment in either the California supply or the competing supply from nearby Florida.

Particularly relevant are the comparative magnitudes of the $\hat{\lambda}_i$. In each of the markets studied, except Atlanta, $\hat{\lambda}_1 > \hat{\lambda}_2$. Thus, each of the five major terminals was more often characterized by relative undersupply than oversupply of celery, and no support is found for the Berger et al. conjecture that California causes market gluts in the major eastern markets.

The estimates of mean arbitrage costs in table 1 are all highly significant. With one exception, Chicago, \hat{T}_C^i is increasing in distance from the producing area. The Chicago result may signal relative inefficiencies or imperfect competition in the Chicago-area transportation and marketing network. Also interesting is the significant portion of wholesale California celery prices that is due to shipment costs. This figure ranges from 28% in Los Angeles to 52% in Atlanta, based on mean 1985–88 wholesale prices.

Transportation costs typically account for a major portion of total transactions costs. Because of the assumption of constant mean trans-

actions costs in the model (see footnote 2), possible trends in transportation costs over the period of analysis are of concern. Monthly shipping-point-to-terminal-market truck transportation costs per crate of celery for the 1985–88 period of analysis were obtained from the U.S. Department of Agriculture, Agricultural Marketing Service (USDA AMS), but observations were available for only Atlanta, Chicago, and New York. These data were tested for trends in transportation costs during the period; no statistically significant time trend was detected in these per crate charges. This result supports the assumption that mean transactions costs were constant over this four-year period.¹²

As noted, celery marketing exhibits a seasonal pattern coinciding with Florida's winter-spring production cycle versus summer-fall marketing when California competes with a number of small northern suppliers. Also, truck shipment rates often exhibit a seasonal pattern, with higher rates during summertime peak demand periods. For these reasons the arbitrage model defined in (9) was estimated separately for the December–May and June–November periods. The results are contained in table 2.¹³

No seasonal pattern in the probability of efficient arbitrage is evident in the Los Angeles and San Francisco markets. This result is not

¹² A second analysis involved subtraction of transportation costs from the terminal market price and estimation of the switching regression model for Chicago and New York. Except for the anticipated differences in the magnitude of T , the probabilities of binding arbitrage were essentially unchanged: 0.15 in Chicago and 0.39 in New York versus 0.19 and 0.36, respectively, in table 1.

¹³ Seasonal models were not run for Atlanta because of limited observations. No California celery price was recorded in Atlanta for many weeks because of insufficient shipments.

Table 2. Seasonal Parameter Estimates for California Celery Arbitrage

	Los Angeles		San Francisco		Chicago		Boston		New York	
	Winter	Summer	Winter	Summer	Winter	Summer	Winter	Summer	Winter	Summer
T	2.55 (19.52) ^a	2.07 228.39	3.68 (2.10)	3.98 (25.28)	5.65 (26.37)	5.13 (18.61)	4.91 (19.69)	4.88 (5.95)	4.46 (16.38)	4.65 (7.72)
σ^2	0.28 (2.25)	0.51 (6.46)	1.14 (3.20)	1.32 (3.43)	0.28 (1.96)	0.15 (2.79)	0.59 (1.68)	0.97 (1.01)	0.87 (3.49)	1.35 (3.64)
σ^2_{λ}	3.71 (3.42)	1.57 (1.99)	9.56 (2.08)	5.38 (1.23)	3.04 (4.84)	2.45 (4.06)	3.74 (3.43)	2.59 (2.61)	4.34 (3.33)	2.67 (2.08)
λ_1	0.14 (2.51)	0.49 (3.84)	0.18 (1.87)	0.06 (0.78)	0.32 (3.29)	0.76 (6.14)	0.42 (2.32)	0.97 (1.61)	0.43 (2.39)	0.80 (8.66)
λ_2	0.35 (3.14)	0.00 ^b	0.04 (0.91)	0.08 (0.92)	0.38 (3.08)	0.14 (2.18)	0.10 (1.34)	0.00 (0.00)	0.06 (1.08)	0.00 ^b
$1 - \lambda_1 - \lambda_2$	0.51	0.51	0.78	0.86	0.30	0.10	0.48	0.03	0.51	0.20
Log likelihood	-163.83	-153.13	-187.47	-190.03	-181.78	-182.22	-177.93	-185.32	-185.01	-201.59
Observations	100	109	100	109	100	109	100	109	100	109
T as % of mean price	30	25	37	40	48	43	42	41	39	41

^a t -statistics are in parentheses.^b Constrained to zero in estimation.

surprising because these markets are served exclusively by California supplies and, hence, are relatively insulated from changes in the seasonal production pattern. In contrast, a seasonal pattern does emerge in the major midwestern and eastern terminals. In all three cases the estimated probability of efficient arbitrage is greater in the winter than summer and in all cases λ_1 (λ_2) is higher in the summer (winter).¹⁴

These results indicate that it is more difficult for California shippers to engage in efficient arbitrage in the summer-fall months than in the winter and spring. This outcome may well be caused by unpredictable patterns of supply by the small, northern celery suppliers during these months. Efficient arbitrage during winter-spring months is likely simplified by the presence of a single main competitor, Florida, with a coordinated marketing mechanism. In particular, Florida supplies likely are easier to predict and monitor than supplies from the small northern producers, and, as noted, through the substitute relationship, Florida shipments may be quite effective in eliminating gluts or shortages for California celery.

Also interesting to note is that the estimated transactions costs, \hat{T}_C^i , are not consistently higher in the summer and fall despite generally higher truck shipment charges during these months. In fact, in three of the five cases the estimated winter-spring margin was higher, indicating that other components of the margin rise in the winter and spring to offset lower truck rates.

Florida Celery

The results from estimation of the arbitrage model for Florida celery in table 3 must be interpreted cautiously because of the market power considerations discussed above. The probabilities $1 - \hat{\lambda}_1 - \hat{\lambda}_2$ are very similar for all of the Florida markets analyzed, ranging from 0.35 to 0.44. However, it is difficult to provide a firm interpretation of this result. If Florida acted as a competitor in these markets and engaged in FOB pricing, the relatively low values for $1 - \hat{\lambda}_1 - \hat{\lambda}_2$ would indicate a failure on Florida's part to

¹⁴ Likelihood ratio tests were used to test the hypothesis of no seasonal structural change. The values of the test statistic were San Francisco: 3.54, Chicago: 6.82, and Boston: 19.96. The pooled model in table 1 imposes 5 restrictions and $\chi^2_{0.05} = 11.1$. Thus, we reject formally the hypothesis of no seasonal structural change for Boston but not for Chicago or San Francisco. (Similar tests were not performed for Los Angeles and New York because in each case the nonnegativity constraint $\lambda_i \geq 0$ was binding for λ_2 in the summer-fall model.)

Table 3. Parameter Estimates for Florida Celery Arbitrage

	Chicago	Boston	New York	Atlanta
T	4.58 (14.79) ^a	2.33 (14.45)	1.93 (8.84)	2.76 (19.74)
σ_v^2	0.37 (2.01)	0.30 (2.59)	0.42 (1.24)	0.16 (1.89)
σ_u^2	2.89 (2.86)	2.10 (3.83)	2.83 (2.98)	1.89 (4.57)
λ_1	0.06 (0.98)	0.23 (2.21)	0.49 (3.35)	0.16 (2.61)
λ_2	0.57 (2.34)	0.33 (2.94)	0.07 (0.57)	0.49 (3.82)
$1 - \lambda_1 - \lambda_2$	0.37	0.44	0.44	0.35
Log likelihood	-81.53	-173.21	-99.21	-170.82
Observations	48	111	61	116
T as % of mean price	41	25	20	29

^a t -statistics are in parentheses.

engage in efficient arbitrage despite its coordinated marketing mechanism. However, the result could also reflect dynamics in the price discrimination equilibrium (10). The price mark-up term, M^i , in equation (11) may change from period to period in response to variations in Florida's residual demand elasticity caused by changes in competing supplies.

Florida is a minor supplier in the Chicago celery market, with an average 2.4% share over 1985–88. The T^i for Chicago in table 3 is comparatively high, as was true for California shipments. Florida is a more significant player in the Boston and New York markets with 26.8% and 17.8%, respectively, of total fresh market sales during the 1985–1988 period of analysis. The T^i for these geographically proximate terminals are similar, \$2.33 and \$1.93 for Boston and New York, respectively.

The most interesting facet of the Florida analysis is comparison of the Boston, New York, and Atlanta markets. Florida is the dominant seller in Atlanta during the winter. In fact, during some weeks, few, if any, California shipments occur. Although Boston (New York) is roughly 850 (650) miles farther from the Florida producing regions than Atlanta, our estimates of the T_F^i indicate that the margin per crate was actually \$0.43 less for Boston and \$0.83 less for New York than Atlanta. This result is evidence that Florida is employing a discriminating spatial pricing system of the type described previously. Per crate truck shipment rates for celery gathered by the USDA AMS further document this result. These rates were not available for Boston, but the average over 1985–88 was \$2.54 for New York and \$1.06 for Atlanta. Truck shipment rates represent only a portion of the

transactions costs involved in marketing celery. They clearly suggest, however, a significant markup above competitive FOB prices in Atlanta and a healthy amount of freight absorption in New York, just as the relevant spatial theory (Greenhut, Norman, and Hung; Faminow and Benson) predicts.

Some Dynamics

Because of the concerns about short- versus long-run equilibrium in arbitrage (Ravallion, Goodwin and Schroeder, Goodwin, Grennes, and Wohlgenant), the extended SH arbitrage model for California celery was reestimated for alternative formulations of the arbitrage condition. Whereas the prototype model in (8) assumes that integration between contemporaneous FOB and terminal prices is the relevant equilibrium concept, alternative formulations may be based on lagged price relationships as suggested by Goodwin, Grennes, and Wohleganant.

A price determination process whereby the law of one price may emerge for lagged prices can be motivated as follows: California FOB prices are based on current supply and demand conditions facing shippers. This demand, in turn, is derived from retailers' demands and demands in the terminals markets. No single terminal is apt to be large enough to have a significant influence on the FOB prices. Weekly terminal wholesale prices are based on supply-demand conditions in that market. When, for example, $P_t^i > p_{C,t} + T_t^i$, a profit opportunity for shipments to market i is signalled if shipments occur instantaneously. The same market signal is received when shipments occur only with a time

lag if the contemporaneous price is used as the expected future price, i.e., $E_t[P_{t+1}^i] = P_t^i$.¹⁵

If shipments arrive with, say, a j period lag, the relevant integration condition is between $p_{C,t}$ and P_{t+j}^i . Given the speed of transcontinental truck shipments, the longest feasible lag to consider for celery is one week. To test for lagged price integration, the extended SH arbitrage model was reformulated and estimated by alternatively substituting P_{t+1}^i or $(P_t^i + P_{t+1}^i)/2$ for P_t^i everywhere in (8) and redefining Y_t accordingly in specifying the f_t^i .

Results from estimating these models are provided in table 4 where the models based on P_{t+1}^i and $(P_t^i + P_{t+1}^i)/2$ are denoted as the "lagged" model and "average" model, respectively. Results for the prototype model (8) (table 1) based on contemporaneous prices are included for comparison.¹⁶

The table shows that the probability, $1 - \lambda_1 - \lambda_2$, of observing binding arbitrage in the Los Angeles and San Francisco terminals is similar across the three arbitrage conditions. This result is again consistent with their proximity to the production region. However, the probability of observing the lagged arbitrage regime, $P_{t+1}^i - p_{C,t} = T_t^i + v_t^i$, is greater in each of the eastern and midwestern markets than is the probability of observing contemporaneous arbitrage, $P_t^i - p_{C,t} = T_t^i + v_t^i$. This result provides some support for the hypothesis of lagged adjustment to market conditions as distance from the producing to the consuming region increases. The probability of observing the "average" relationship, $(P_t^i + P_{t+1}^i)/2 - p_{C,t} = T_t^i + v_t^i$, is even greater in these three terminals.

Further information on the dynamics of price response in this market is obtained by comparing the log likelihood values in the table. For this purpose the contemporaneous integration model was reestimated to reflect observations lost in the lagging process ($n = 207$ vs. $n = 209$). This likelihood value is denoted as the "comparison likelihood" in the table.¹⁷ The log likelihood is greatest for the contemporaneous integration model for both Los Angeles and San Francisco, but the model using the average of P_t^i and P_{t+1}^i generates the greatest log likelihood

Table 4. Comparison of Alternative Arbitrage Models for California Celery

	Los Angeles			San Francisco			Chicago			Boston			New York		
	Contemporaneous	Average	Lagged	Contemporaneous	Average	Lagged	Contemporaneous	Average	Lagged	Contemporaneous	Average	Lagged	Contemporaneous	Average	Lagged
T	2.35 (28.28)	2.28 (14.86)	2.13 (8.29)	3.83 (32.66)	3.81 (28.02)	3.81 (27.09)	5.45 (1.86)	5.82 (47.41)	5.65 (41.15)	4.84 (5.82)	5.64 (55.53)	5.56 (3.69)	4.53 (11.02)	4.70 (18.06)	4.40 (18.27)
σ_v^2	0.62 (5.28)	0.68 (3.14)	1.93 (2.81)	1.28 (5.39)	1.46 (5.36)	2.55 (6.62)	0.33 (1.68)	1.41 (5.51)	1.18 (6.23)	0.87 (2.14)	1.58 (5.06)	1.74 (6.12)	1.05 (3.24)	0.99 (1.27)	1.27 (4.31)
σ_u^2	3.94 (3.52)	3.71 (3.44)	7.74 (2.33)	7.97 (2.29)	7.81 (2.05)	17.33 (2.06)	2.23 (5.61)	5.53 (1.33)	7.57 (2.80)	2.65 (2.42)	3.88 (1.20)	9.33 (2.02)	3.53 (4.46)	2.39 (3.29)	5.66 (4.89)
λ_1	0.15 (2.27)	0.25 (1.96)	0.22 (1.33)	0.12 (1.80)	0.12 (1.40)	0.09 (1.59)	0.55 (3.69)	0.05 (0.70)	0.16 (1.95)	0.78 (1.69)	0.08 (0.83)	0.13 (1.61)	0.61 (2.46)	0.45 (2.23)	0.55 (3.85)
λ_2	0.09 (16.12)	0.12 (2.14)	0.06 (0.98)	0.05 (1.27)	0.03 (0.94)	0.04 (1.22)	0.26 (2.60)	0.05 (0.95)	0.07 (1.79)	0.04 (0.63)	0.00 (0.00)	0.02 (0.88)	0.03 (0.95)	0.00 (0.00)	0.02 (0.87)
$1 - \lambda_1 - \lambda_2$	0.76 (323.44)	0.63 (351.98)	0.72 (427.94)	0.83 (379.47)	0.85 (380.13)	0.89 (437.20)	0.19 (367.42)	0.90 (357.60)	0.77 (386.16)	0.28 (373.22)	0.92 (358.26)	0.85 (395.77)	0.36 (391.90)	0.55 (363.15)	0.43 (415.41)
Log likelihood Observations	209	207	207	209	207	207	209	207	207	209	207	207	209	207	207
Comparison likelihood ($n = 207$)	-320.64	-320.64	-320.64	-376.21	-376.21	-376.21	-364.65	-364.65	-364.65	-370.15	-370.15	-370.15	-386.00	-386.00	-386.00

¹⁵ This formulation is merely the condition that celery prices are martingales (Alchian), i.e., the mathematical expectation of the future price is the current price.

¹⁶ Atlanta was not included in this analysis because the periodic absence of California prices made it difficult to handle the lags.

¹⁷ Results for the contemporaneous integration model for $n = 209$ and $n = 207$ are nearly identical; hence, only the $n = 209$ results are reported.

Table 5. Parameter Estimates for Price Premium Model

	Chicago	Boston	Atlanta	New York
α	0.76 (1.24) ^a	2.90 (4.70)	2.57 (2.84)	2.31 (9.42)
σ_μ^2	0.34 (0.42)	0.47 (1.21)	0.56 (0.61)	0.04 (0.09)
σ_u^2	1.66 (2.04)	1.53 (2.04)	2.01 (2.03)	3.65 (7.80)
λ_1	0.71 (1.84)	0.39 (1.56)	0.57 (1.18)	0.45 (4.61)
λ_2	0.29 (1.12)	0.60 (2.07)	0.42 (1.13)	0.55 (4.84)
$1 - \lambda_1 - \lambda_2$	0.00	0.01	0.01	0.00
Log likelihood	-89.63	-208.48	-124.56	-195.76
Observations	53	119	66	63

^a *t*-statistics are in parentheses.

in the midwestern and eastern terminals. These results thus provide support for an arbitrage model based on contemporaneous prices in terminals near the producing area and on a short lagged price relationship elsewhere.

Price Premia Results

The final aspect of the empirical analysis involved direct comparison of California and Florida celery prices in the various terminal markets. The results from estimation of the price premium model (13) are provided in table 5. The estimated price premia, $\hat{\alpha}^i$, are similar for the Atlanta, Boston, and New York markets, ranging from \$2.31 to \$2.90 per crate. The $\hat{\alpha}^i$ is much lower, \$0.76, in Chicago, but its *t*-statistic indicates that this coefficient was estimated imprecisely.

The most interesting facet of this analysis is that the estimated probability, $1 - \hat{\lambda}_1 - \hat{\lambda}_2$, of observing a stable price premium is nearly zero in each of the markets. It bears repeating that if buyers viewed California and Florida celery as perfect substitutes subject to a price premium, then buyer arbitrage should ensure that prices diverge by exactly the premium whenever product from both states is available. Thus, the conclusion is that buyers do not view California and Florida celery as perfect substitutes, and their price difference is free to vary somewhat above or below the estimated premium in response to market conditions.

Conclusions

This paper has extended Spiller and Huang's methodology to test market integration and ap-

plied it in an agricultural markets setting. One extension was to regions known to be linked in trade, specifically the common situation in agriculture of one or a few localized production areas. When perfect competition in product allocation and FOB pricing can be safely assumed, this extended model allows for tests of three regimes: (a) efficient arbitrage (i.e., the law of one price), (b) relative shortage, and (c) relative glut. A second extension enabled limited testing of products' substitutability in demand.

Results from application of the extended models to celery marketing indicated that California shipments to out-of-state markets in nearly all cases departed with significant probability from the efficient arbitrage condition. However, no evidence was found to support the hypothesized eastern market glut scenario. The probability of binding arbitrage did increase substantially with the introduction of lags associated with physical transfer of the product from California to distant markets. A seasonal pattern in California marketing efficiency was detected in midwestern and eastern terminals, with efficient arbitrage apparently made easier in winter-spring months by the presence of a single main competitor, Florida, versus many small northern rivals in the summer-fall months.

Interpretation of results from analysis of Florida celery was difficult because of concerns that Florida may practice spatial price discrimination. Comparison of results for the various terminals tended to support the position that Florida charges prices well above the FOB level in the nearby Atlanta market and absorbs some portion of freight charges in New York. The Florida results highlight that the failure to observe binding arbitrage across pairs of markets

may be due to autarkic markets, inefficient arbitrage, or imperfect competition. Unless it is possible to rule out some of these considerations, it will be very difficult to offer a single explanation for a failure to observe the law of one price.

The methodology illustrated in this paper is parsimonious in its data requirements, needing only time-series price data for alternative cities, regions, countries, or product forms. Yet, by exploiting the flexibility of the approach and the available economic theory concerning the underlying market structure, evidence may be generated on a number of market parameters including market integration, arbitrage efficiency, magnitude of marketing margins, product substitutability, and competitiveness of markets.

In many cases more detailed information will be needed than can be provided by an approach that relies solely on price data. Of course, the downside to detailed structural market models is their demands on data, which often compel undesirable levels of aggregation and use of dubious proxies for otherwise unmeasurable variables.

[Received May 1990; final revision received October 1990.]

References

- Aigner, D., C. A. K. Lovell, and P. Schmidt. "Formulation and Estimation of Stochastic Frontier Production Function Models." *J. Econometrics* 6(1977):21-37.
- Alchian, A. A. "Information, Martingales, and Prices." *Swedish J. Econ.* 76(1974):3-11.
- Baker, J. B., and T. F. Bresnahan. "The Gains from Merger or Collusion in Product-Differentiated Industries." *J. Indust. Econ.* 33(1985):427-44.
- Berger, L., T. Hebert, D. Ricks, and J. Shaffer. "The U.S. Celery Production and Marketing Subsector: An Evaluation of Market Coordination and Performance." Dep. Agr. Econ. Agr. Econ. Rep. No. 517, Michigan State University Oct. 1988.
- Buccola, S. T. "Pricing Efficiency in Centralized and Noncentralized Markets." *Amer. J. Agr. Econ.* 67(1985):583-90.
- . "Risk Preferences and Short-Run Pricing Efficiency." *Amer. J. Agr. Econ.* 65(1983):587-91.
- Carter, C. A., and N. A. Hamilton. "Wheat Inputs and the Law of One Price." *Agribus*. 5(1989):489-96.
- Faminow, M. D., and B. L. Benson. "Integration of Spatial Markets." *Amer. J. Agr. Econ.* 72(1990):49-62.
- Goodwin, B. K., and T. C. Schroeder. "Testing Perfect Spatial Market Integration: An Application to Regional U.S. Cattle Markets." *N. Cent. J. Agr. Econ.* 12(1990):173-86.
- Goodwin, B. K., T. J. Grennes, and M. K. Wohlgenant. "A Revised Test of the Law of One Price Using Rational Price Expectations." *Amer. J. Agr. Econ.* 72(1990):682-93.
- Greenhut, M. L., G. Norman, and C. S. Hung. *The Economics of Imperfect Competition: A Spatial Approach*. Cambridge: Cambridge University Press, 1987.
- Hayes, D. J., T. I. Wahl, and G. W. Williams. "Testing Restrictions on a Model of Japanese Meat Demand." *Amer. J. Agr. Econ.* 72(1990):556-66.
- Horowitz, I. "Market Definition in Antitrust Analysis: A Regression-Based Approach." *S. Econ. J.* 48(1981):1-16.
- Kilmer, R. L. "A Review of the Current Pricing Systems of Celery, Sweet Corn, and Potatoes." Food and Resource Econ. Dep. Econ. Info. Rep. No. 171, University of Florida, Nov. 1982.
- Ravallion, M. "Testing Market Integration." *Amer. J. Agr. Econ.* 68(1986):102-9.
- Richardson, D. J. "Some Empirical Evidence on Commodity Arbitrage and the Law of One Price." *J. Int. Econ.* 8(1973):341-51.
- Spiller, P. T., and C. J. Huang. "On the Extent of The Market: Wholesale Gasoline in the Northeastern United States." *J. Indust. Econ.* 34(1986):131-45.
- Stigler, G. J., and R. A. Sherwin. "The Extent of the Market." *J. Law and Econ.* 37(1985):555-85.
- Takayama, T., and G. G. Judge. *Spatial and Temporal Price and Allocation Models*. Amsterdam: North-Holland Publishing Co., 1971.
- Taylor, T. G., and R. L. Kilmer. "An Analysis of Market Structure and Pricing in the Florida Celery Industry." *S. J. Agr. Econ.* 20(1988):35-43.

An Empirical Analysis of Economic Performance Under the Marketing Order for Raisins

Ben C. French and Carole Frank Nuckton

This study utilizes a dynamic econometric model of the California raisin industry to compare predictions of prices, production, profits, and related measures under the volume control program with predictions under several no-control scenarios. The outcomes are evaluated in relation to performance criteria proposed by a USDA study team. The twenty-two-year comparisons suggest that the public interest may have been well served by the raisin volume control program, or at worst there was no significant welfare loss.

Key words: econometric model, marketing order, raisins.

Marketing order programs that control the quantity of a product marketed have been authorized for a number of fruit and vegetable commodities as a means of mitigating the negative effects of variable supply and demand on returns to producers. Although there have been some theoretical analyses of the price stabilization aspects of such controls and a few empirical studies, views on applicable performance criteria differ and empirical measurement of social welfare indicators has proved quite difficult (Heifner et al.; Armbruster and Jesse; U.S. GAO; Jesse; French 1982, 1988).

The California raisin industry has had a federally authorized volume control program in effect since 1949. Under this program, the industry is permitted to set aside a portion of the crop in a reserve pool. Raisins in the reserve pool may be returned to commercial sales when market conditions are favorable or disposed of in noncompetitive outlets such as government purchases for school lunches, sales for alcohol manufacture and cattle feed, and donations to charity. Some reserve pool raisins have been exported at prices competitive with world prices but below those on the domestic market. Significant quantities of raisins have been allocated

to reserve pools in all years except when supply was low relative to demand.

The objective of this study is to provide evidence as to possible differences in the behavior of prices, profits, production, and related measures of raisin industry performance in the absence of the volume control program. If the volume control provisions of the marketing order were terminated, grower and consumer prices would be expected to be more variable over time and lower in the short run. Some of the burden of inventories likely would transfer to packers. However, with delayed production response feeding back through plantings and removals, the pattern of outcomes over a period of years is not obvious, hence the need for dynamic empirical analysis. While the findings of the study are specific to the raisin marketing order, the modeling approach may be of general interest, and the conclusions may be relevant to the evaluation of similar types of controls for other commodities.

The study procedure is to utilize an econometric model of the raisin industry system to generate and compare historical predictions of prices, outputs, and returns under the control program with predictions of the model modified to reflect likely conditions without volume controls in effect.¹ The approach is a modification

Ben C. French is a professor, Department of Agricultural Economics, University of California, Davis, and Carole Frank Nuckton is assistant to the director, University of California Agricultural Issues Center, Davis.

Giannini Foundation Research Paper No. 956.

The authors are appreciative of helpful comments by Gordon A. King and two anonymous *Journal* reviewers.

¹ The historical approach has an advantage over a forward analysis in that random elements such as yields and estimated stochastic disturbances are known. In a forward analysis, yield fluctuations and other disturbances would have to be generated from the variance-covariance structure of the model by repeated runs.

and extension of previous studies of the effects of eliminating marketing order controls affecting annual volumes of fresh lemons (French and Bressler; Kinney et al.), fresh oranges (Thor and Jesse), and cling peaches (Minami, French, and King). A problem encountered in all of these studies was that because the control provisions were in effect throughout the period of available data, possible differences in equation parameters without the marketing order could not be determined (the well-known Lucas critique). Although the problem was recognized, none of the studies attempted to overcome it, and Kinney et al. made only qualitative evaluations.

A similar problem was encountered in studying the raisin industry's volume control program. The method of dealing with the Lucas criticism was to specify several alternative scenarios which involve plausible changes in equation parameters without the reserve pool provision of the raisin marketing order. For example, a study by French and King (FK 1988) of the effects on planting response of terminating the cling peach volume program in 1972 provides a basis for specifying possible changes in the raisin grape planting response.

In comparing economic performance under the volume control program with performance under the no-control scenarios, all exogenous variables and equation parameters remain the same except for those associated with elimination of volume control. To take account of the random fluctuations in supply and demand which motivated the establishment of the reserve pool program, the estimated disturbances of the stochastic equations and actual yield variations are retained as exogenous variables in the various simulation runs.

Because of difficulties in developing meaningful empirical measures of conventional social welfare indicators, the findings are evaluated in relation to six of seven performance criteria proposed by an independent marketing order study team appointed by the U.S. Department of Agriculture (Polopolus et al.). To meet public interest requirements, the study team proposed that a volume control program should not (a) permit farmers to earn persistent above normal profits, (b) increase price variability and uncertainty, (c) impose disproportionate burdens on particular classes of growers or handlers, (d) contribute to chronic surpluses, (e) result in a waste of resources, (f) reduce net revenues to producers, and (g) result in consumer prices persistently higher than justified by costs including a normal profit. Results of our analysis allow evaluation

of the raisin marketing order volume control provision under (a), (b), and (d)–(g) of these criteria.

The Historical Model

Most raisin-grape production is from Thompson seedless grapes and most of the raisin crop is sun dried and is referred to as natural Thompson seedless raisins (NTS). The econometric model of the NTS industry is a slightly revised version of one by Nuckton, French, and King (NFK).² The model groups the economic activities of the industry into four blocks: (1) growers' raisin-grape vine planting and removal decisions and the resulting bearing acreage and raisin-grape production; (2) growers' allocation of raisin-grape production between drying and crushing for wine; (3) the Raisin Administrative Committee's (RAC) decision rules allocating NTS production between free and reserve tonnage and the determination of the outcome of bargaining between independent packers (not associated with the large grower cooperative) and the growers' Raisin Bargaining Association (RBA) for the field price for free raisin tonnage; and (4) domestic and foreign demand functions, packer pricing and inventory decisions, and the RAC's establishment of an export price for reserve pool quantities. The model is simultaneous within blocks 2, 3, and 4; but, since block solutions are determined sequentially within a crop year, it is recursive among blocks.

Model equations were estimated with data from 1963–83. Out-of-sample projections for 1984 and 1985, reported in NFK, were generally within 95% confidence intervals except for predictions of domestic bulk raisin sales (*QB*) and sales to the United Kingdom and Scandinavia, which were underpredicted. The deviations between actual and predicted values were entered exogenously in the simulations so any structural change that may have occurred in 1985 and 1985 would have little effect on the comparative simulations. A brief explanation of the model specifications, estimation procedures, and empirical results follows.

² The main revisions are in the raisin-grape planting and removal equations where an improved measure of grower returns is used and in the reported beginning stock equation where a revised measure of total product movement is used. Also, some typographical errors in table 3 of NFK (not contained in their simulations) were corrected. The NFK simulations explored the effects of the export incentive program and changes in exchange rates, tariffs, production costs, and population. They did not consider the effects of terminating the volume control program.

Block 1. Production of Raisin-Type Grapes

Annual production of raisin-type grapes is determined by bearing acres and yield, with yield treated as an exogenous variable. Bearing acres in t are determined by the identity, $BA_t = BA_{t-1} + PL_{t-3} - R_{t-1}$ where BA is bearing acres, PL is plantings, and R is acres removed.³

The model for estimating planting and removal functions was adapted from French, King, and Minami and FK 1988. Plantings are a function of existing acreage, less any removals, and grower projections of net returns over the expected life of the vines. Expected long-run returns are modeled as a function of past prices less unit costs. Young acreage, a significant factor affecting future industry supply expectations for peaches, was not significant for raisins because the productive life of grape vines far exceeds that of peach trees. Measures of variability of past profits were not significant predictors in the plantings equations for either raisin grapes or peaches.

Acreage removed from production each year is determined mainly by declining productivity associated with biological factors but is also affected by growers' expected returns. If growers expect favorable returns the next year, they may retain some marginal vineyards another year, decreasing removals, and vice versa. Profit expectations for year t are related to recent profit experience; however, the expectations horizon is shorter than for planting decisions.

Ordinary least square estimates of the planting and removal functions are (values in parenthesis are t -statistics; d is the Durbin-Watson statistic):

plantings ($R^2 = .83$, $d = 1.78$)

$$(1) \quad PL_t / (TA - R)_{t-1} = .01211 + .000979 \text{ NGD3}_{t-1}, \quad (7.59) \quad (9.55)$$

removals ($R^2 = .29$, $d = 1.59$)

$$(2) \quad R_t / BA_{t-1} = .01702 - .000329 \text{ NGD2}_{t-1}, \quad (8.50) \quad (-2.79)$$

where TA is total acres and NGD is grower net return for raisin-type grapes dried and crushed (prices less unit costs), deflated by the gross national product deflator ($GNPD$). Averages over

three-year ($NGD3$) and two-year ($NGD2$) periods were the best predictors among alternative lag specifications. The low R^2 value for the removal equation and the high significance of the intercept reflect the dominance of nonprice factors in removal decisions.

Block 2. Allocation of Raisin-Type Grapes

Raisin grapes are utilized for drying (1963–85 average = 52.8%), crushing (34.9%), canning (2.5%), and selling fresh (9.8%). Cultural practices for fresh market grapes restrict switching between raisins and fresh sales in the short run and the minor sales for canning could not be modeled separately. Fresh market production (QF) and sales for canning ($QCAN$) were treated as exogenous with the balance of production (Q) allocated endogenously between drying (QR) and crushing (QC).

At the time growers make their allocation decisions, the demand for raisin grapes to crush is known but that for raisins is not. In a purely competitive environment, growers would allocate Q so as to achieve equivalent expected net returns in both outlets, given the observed demand for grapes crushed and the projected demand for raisins. The static competitive (reduced-form) solution would express QR as a function of Q , the demand shifters for grapes crushed, and the expected values of demand shifters for raisins. However, because of the reserve pool operations of the marketing order and the bargaining activities of the RBA , prices received by growers for grapes dried ($PGRD$) have often exceeded the crush price ($PGCD$) by substantially more than would be expected based on differences in costs. Hence, the competitive allocation model seems inappropriate.

An alternative procedure is to view the allocation process as a conditional supply relationship. The quantity of raisin grapes converted to raisins (QR) is expressed as a function of Q , net return to crush, expected net return to grapes dried, and a lagged value of QR reflecting a partial adjustment process.

In forming their expectations of returns from raisin production, growers are assumed to be forward looking, giving consideration to expected supply and market conditions. However, because of difficulties in forecasting the outcomes of RAC and RBA actions and domestic and world demand for raisins, the dominant factors affecting expected net returns for raisins appear to be the size of the raisin grape crop avail-

³ A very small proportion of plantings may be removed before reaching bearing age, but for simplicity this is assumed to be zero.

able for crushing and drying (Q) and recent profit experience. For the latter, a three-year average of past net grower returns (deflated) per ton of grapes dried ($RRD3$) was the best predictor among alternatives considered.

The demand for raisin grapes crushed is influenced by a number of factors not of direct interest in this study, other than to account for their aggregate effects on the price of raisin grapes crushed. Therefore, our estimate of the derivative of the grower crush price ($PGCD$) with respect to quantity of raisin grapes crushed (QC) was obtained from an econometric study of the wine industry by Wohlgenant. The variations in the wine price unexplained by the Wohlgenant coefficient enter the model as exogenous disturbances.

To account for simultaneity, the allocation-to-dry equation was estimated by instrumental variables where the instruments are the predetermined variables Q_t , $RRD3_{t-1}$, QC_{t-1} , QR_{t-1} , $GCWD_t$, $PGCD_{t-1}$, and a constant. $GCWD$ is a measure of deflated cost per ton to produce raisin grapes for crush, and other variables are as defined above. The empirically estimated allocation system is (t -statistics in parentheses) as follows:

Raisin allocation ($R^2 = .89$, Durbin $h = .638$)

$$(3) \quad QR_t = -444,439 + .6037 Q_t + 2178.9 RRD3_{t-1} - 4791.4 (PGCD - GCWD)_t + 0.2623 QR_{t-1} \\ (-1.60) \quad (4.45) \quad (1.81) \quad (-1.99) \quad (2.06)$$

Crush demand

$$(4) \quad PGCD_t = PGCD_{t-1} - .000045(QC_t - QC_{t-1}).$$

Crush allocation

$$(i) \quad QC = Q - QR.$$

Block 3. Allocation and Pricing of NTS Deliveries

The quantity of NTS raisins delivered (DEL) is obtained by multiplying QR , determined in block 2, by a conversion factor. The RAC allocates DEL between "free" tonnage (QFR) and a reserve pool (RES).⁴ In short-crop years, all production may be declared free. The marketing or-

der lists several factors to be considered by the RAC in establishing QFR (NFK, p. 13). Significant predictors of the RAC free tonnage decisions are, on the supply side, deliveries (DEL) plus stocks carried in from the previous year in the reserve pool (CI) and in the hands of packers (SP); on the demand side, previous year packer price for raisins (PD), domestic movement (QD), and a variable (X) to account for a shift in the perception of the export market beginning in 1977. Before 1977, raisin exports were regarded primarily as distress sales from the reserve pool (except in short-crop years). With the strengthening of export markets in the mid-1970s, the RAC shifted its policy to view exports as free tonnage to be sold in regular channels. This shift is modeled by assigning X a value of zero before 1977; the value of lagged total exports (QXT_{t-1}) was used thereafter.

The price packers pay growers for free tonnage is determined by the bargaining process between independent (non-co-op) packers and the grower RBA. The bargaining structure suggests that a competitive farm-level demand function may not be defined and unique market equilibrium solutions may not exist. However, following French (1987), consistent predictions of the (deflated) price outcomes of the bargaining process (PFD) are obtained with a func-

tion that includes essentially the same explanatory variables as would be contained in a competitive grower-level raw product demand function: (a) the free tonnage quantity (QFR) and (b) expected packer net returns in the forthcoming market year, determined primarily by lagged FOB prices and costs, total carryover stocks, and the variable X to account for the policy shift about exports.

Because of (mostly small) inconsistencies in the reports of packer carryin stocks (SP), the identity which relates actual carryin stocks to the previous year's beginning stocks plus quantity packed less sales is replaced by a stochastic equation in the same variables. This equation predicts reported free carryin stocks in the hands of packers. Reported stocks enter the information set in place of actual but unobserved stocks.

The stochastic equations plus key identities which complete block 3 are as follows (t -statistics in parentheses):

Reported stocks ($R^2 = .97$, Durbin $h = -.265$)

⁴ Besides the initial RAC free tonnage declaration, packers may purchase additional tonnage from the reserve pool as the marketing year progresses, for which they pay the free tonnage price plus interest and storage charges. The variable QFR is defined to include both the RAC initial declaration and additional purchases.

$$(5) \quad SP_t = -5873.1 + .904 QFR_{t-1} + .923 SP_{t-1} - .848 (QP + QB + QXF)_{t-1} \\ (-1.30) \quad (11.85) \quad (13.21) \quad (7.58)$$

Free tonnage ($R^2 = .85$, $d = 2.02$)

$$(6) \quad QFR_t = -34,531 + 39.84 PD_{t-1} + .78 (QP_t + QB)_{t-1} - 1.92 SP_t \\ (-.76) \quad (2.98) \quad (2.02) \quad (-4.05) \\ + .40 (DEL + CI)_t + 1.59 X_t \\ (6.38) \quad (3.89)$$

Free tonnage price ($R^2 = .79$, $d = 2.07$)

$$(7) \quad PFD_t = 890.52 + .431 (PD - APCD)_{t-1} - 1.00 (QFR/N)_t \\ (7.77) \quad (4.15) \quad (-4.94) \\ - .19 (SU - QFR)_t/N_t + .0064 X_t \\ (-1.40) \quad (4.67)$$

Reserve carryover

(ii)

$$CI_t = DEL_{t-1} + CI_{t-1} - QFR_{t-1} \\ - QXR_{t-1} - OT_{t-1}$$

Total supply

$$(iii) \quad SU_t = DEL_t + SP_t + CI_t$$

Reserve pool quantity

$$(iv) \quad RES_t = DEL_t + CI_t - QFR_t$$

where variables not otherwise identified are QP and QB , sales in the domestic market (U.S. plus Canada) in package and bulk forms, respectively; QXF , free tonnage exports; QXR , reserve pool exports; PD , the weighted average price for QP and QB sales (deflated by GNPD); $APCD$, the deflated weighted average processing cost per unit for QP and QB ; N , U.S. plus Canadian population; and OT , reserve pool quantity disposed of in other (noncommercial) uses.

Equation (5) was estimated by ordinary least squares. Because the RAC establishes QFR before the grower price (PFD) is bargained for, QFR enters (7) recursively. Equations (6) and (7) were estimated as seemingly unrelated regressions because of likely contemporary correlation of the disturbances.

Block 4. Processed Product Pricing and Demand

Following French and King (1986), packers are viewed as price setters where the FOB price of raisins is set to cover processing plus raw product cost, with further adjustment based on current market conditions, primarily in the domestic market (evidence of such pricing behavior is

provided in NFK, p. 16). Quantities not sold at the offered price are carried to the next season [reported stocks are predicted by equation (5) in block 3].

The price packers set for exports has varied over time according to the export policy adopted by the RAC. In periods of no exports from the reserve pool (1977–80 and short-crop years), the export price was the same as the FOB domestic packer price for bulk raisins (PB) since most exports are in bulk form. During the periods 1963–76 and 1981–85, excluding years with no reserve pool, the RAC sold reserve pool raisins to packers for export at prices such that the weighted average undeflated export price to growers for free and reserve pool exports (PGX) was less than the free tonnage price in nominal dollars (PF); in other years $PGX = PF$. PGX was found to be closely related to the dollar price for Greek raisins in the United Kingdom (an indicator of the world price level for raisins) and the supply of raisins divided by N . The undeflated packer export price (PX) is PGX plus the nominal packer-grower margin for domestic bulk raisins ($PB-PF$).

The price of U.S. raisins in each importing country was computed by adding a transfer cost to PX , then multiplying by the ad valorem duty charged in that country and the U.S. exchange rate with that country, and deflating by that country's CPI. Each foreign market demand function was expressed with per capita sales of U.S. raisins in country I , $QC(I)$, as a function of the deflated price of U.S. raisins in I , the price of competing raisins in I and the per capita deflated income in I .

The price establishment equations, the domestic and foreign demand functions, and the identities and technical relations required to complete the model form a large simultaneous

system in which the predetermined and exogenous variables exceed the number of observations. To deal with this undersize sample problem, the system was divided into two somewhat recursive sub-blocks. Because the FOB price for bulk raisins (PB) is dominantly influenced by variables within the domestic sub-block, it was treated as predetermined with respect to the export sector. The domestic sub-block was estimated by three-stage least squares. In the export block, the PGX equation and the foreign demand equations were estimated as seemingly unrelated regressions. The estimation results are as follows (D appended to price variables indicates deflated value):

Domestic Price Setting and Demand System

Package price ($R^2 = .96$, $d = 1.60$)

$$(8) \quad PPD_t = -458.29 + 1.33 (PCPD + PFD)_t - 1148.3 (QP/SU)_t$$

(5.11) (21.06) (3.99)

Bulk price ($R^2 = .96$, $d = 2.25$)

$$(9) \quad PBD_t = -607.10 + 1.12 (PCBD + PFD)_t + 1620.3 (QP/SU)_t$$

(-7.51) (20.19) (6.17)

Package per capita demand ($R^2 = .72$, $d = 2.28$)

$$(10) \quad (QP/N)_t = 102.93 - .077 PPD_t + 1.161 PSUBD_t - .0047 EUD_t$$

(3.05) (-6.31) (5.17) (-0.43)

Bulk per capita demand ($R^2 = .50$, $d = 1.29$)

$$(11) \quad (QB/N)_t = 289.57 - .073 PBD_t - .530 PCOMD_t + .024 EUD_t$$

(3.78) (-4.32) (-1.00) (1.92)

where variables not otherwise identified are $PCPD$ and $PCBD$, per unit processing costs for raisins in package and bulk form; $PSUBD$ and $PCOMD$, price indexes for substitutes and complements, respectively (see NFK for further development and interpretation); and EUD_t , U.S. per capita deflated personal consumption expenditures.⁵

Export Sector

Grower average export price per ton ($R^2 = .92$, $d = 2.13$)

$$(12) \quad PGX_t = 196.25 + 1.007 PGR_t$$

(1.29) (15.40)

$$- .291 (SU/N)_t$$

(-2.29)

packer export price

$$(v) \quad PX_t = PGX_t + (PB - PF)_t$$

foreign market price ($I = UK, WG, N, S, J$)

$$(vi) \quad PU(I)D_t = [PX_t + TU(I)] \cdot D(I)_t \cdot E(I)_t / CPI(I)_t$$

United Kingdom (UK) per capita demand ($R^2 = .54$, $d = 2.04$)

$$(13) \quad QCUK_t = 525.56 - .318 PUUKD_t$$

(7.22) (-3.11)

$$+ .195 PGUKD_t - .655 EUKD_t$$

(1.03) (-5.30)

West Germany (WG) per capita demand ($R^2 = .59$, $d = 1.37$)

$$(14) \quad QCWG_t = 13.70 - .0168 PUWGD_t$$

(0.78) (3.31)

$$+ .0153 PGWGD_t + .0069 EWGD_t$$

(1.66) (3.57)

Netherlands-Belgium (N) per capita demand ($R^2 = .74$, $d = .83$)

$$(15) \quad QCN_t = 13.06 - .0336 PUND_t$$

(0.65) (-4.58)

$$+ .0204 PGND_t + .0166 END_t$$

(1.65) (6.01)

Sweden-Norway-Denmark (S) per capita demand ($R^2 = .67$, $d = 2.48$)

$$(16) \quad QCS_t = 815.99 - .0786 PUSD_t$$

(6.38) (-3.58)

$$+ .0850 PGSD_t - .0291 ESD_t$$

(2.34) (-2.46)

Japan (J) per capita demand ($R^2 = .67$, $d = 2.48$)

⁵ A variable to reflect the effects of advertising expenditures under the state marketing order and by large packers such as Sun Maid was not included because the effects were thought to remain at about the same general level throughout the period of the data set. The highly successful "dancing raisins" advertising program appeared after the period of analysis.

$$(17) \quad QCJ_t = 252.52 - .00058 \text{ PUJD}_t \\
\begin{matrix} (6.22) & (-6.86) \\ + .0024 \text{ PAJD} - .0001 \text{ EJD}_t, \\ (1.66) & (-1.64) \end{matrix}$$

Total U.S. exports

$$(vii) \quad QXT_t = \sum_I QC(I) \cdot N(I)_t [1 + ROW_t]$$

Variables not previously defined are: PGR , the price of Greek raisins per metric ton (mt) landed in London, nominal dollars; $PG(I)D$, the deflated price per mt of Greek raisins landed in European importing country I , in country I 's currency; $PAJD$, the deflated price of Australian raisins landed in Japan in yen per mt; $TU(I)$, transportation cost from the U.S. to importing country I ; $D(I)$, ad valorem duty plus one, $E(I)$, country I -to-U.S. exchange rate, and $CPI(I)$, country I 's consumer price index; $E(I)D$, deflated per capita income in country I ; and ROW , the ratio (exogenous) of U.S. exports to the rest of the world to exports to the five country groups. The price computed for the Netherlands is used for the Netherlands-Belgium group; the price computed for Sweden, the Sweden-Norway-Denmark group. The negative coefficients for $E(I)D$ in the UK, S and J may reflect a substitution of fresh fruit for dried as incomes have risen.

The equations or rules required to predict the quantity of raisins exported from the reserve pool under the various RAC export policies are not reported here because of space limits and because they are not directly required in the simulations which follow (see NFK, p. 18).

Simulation Analysis

The estimated equations of the four blocks, together with the identities which define the net return variables, form a complete system which, given the values of an initial set of lagged endogenous variables and all values of the exogenous variables, may be solved sequentially for the predicted values of the endogenous variables for all future years. This sequential simulation is used to explore the question, What might have happened if the reserve pool provision of the federal marketing order for raisins had been terminated in 1963, effective in 1964, with all other aspects of the economic environment the same? (1963 was the first year of complete data.) The estimated disturbances of the stochastic equa-

tions are entered as exogenous variables and remain the same with and without the volume control in effect, except as noted. Thus, the sequential solution of the historical model (volume control in effect) generates exact predictions of the observed values of the endogenous variables with which to compare predictions of the no-volume-control scenarios.

Modifications for No-Volume-Control Conditions

If the volume control provision of the federal marketing order were terminated, equation (6), which predicts free tonnage (QFR), would be dropped since there would be no reserve pool and all production would be "free." Equation (12), which predicts the average grower export price (PGX), would be replaced by the identity $PGX = PF$, where PF , is the undeflated free tonnage price. There is little reason to expect that the demand function for raisin grapes for crushing (4), the domestic demand for NTS (10) and (11), or the foreign demand functions, (13) through (17), would differ without volume control. The advertising program under the separate state marketing order (with effects implicit in the estimated demand parameters) is assumed to remain in place with roughly the same general impact. The marketing order does not impose restrictions on packer marketing of free tonnage NTS , so packer FOB pricing practices, (8) and (9), seem unlikely to change significantly; nor would (5) which predicts reported packer carryin stocks. However, four equations—(1), (2), (3) and (7)—might change without the volume control program.

French and King (1988) found that the rate of planting response for a given level of profitability dropped by as much as 45% after the cling peach volume-control marketing order program was terminated. This result suggests a range of alternative scenarios for equation (1). Because French and King found no change in the relation of removal rates to net returns when the peach volume control ended, equation (2) will remain unchanged.

The process of bargaining between packers and the grower bargaining association is assumed to continue without the reserve pool program, with outcomes still predicted by equation (7). However, because all raisin production would be "free," the quantity of NTS produced (DEL) would replace QFR .

If the risk associated with raisin production is

perceived to increase in the absence of the volume control, it is possible that growers would allocate relatively more raisin grapes to the more certain crush market. Therefore, the analysis includes scenarios that tilt the raisin grape allocation equation (3) slightly toward the crush market.

In years of large crops, the RAC set free tonnage (QFR) below DEL , which maintained the free tonnage price of raisins (PFD) generally above equivalent returns per ton of grapes in the crush market ($PGCD$). With a mean ratio of 5.4 short tons of raisin grapes per packed-weight metric ton of raisins, $PFD = 5.4 PGCD$ gives the same gross return per ton for grapes in either outlet. However, 5.4 does not account for the differences in cost and risk factors between the dry and crush outlets. Historically, the actual ratio $PFD/PGCD$ has varied from seven to fifteen.

Without the reserve pool to limit marketable raisin tonnage, PFD as predicted by (7), could be driven to very low levels. However, because growers know the crush demand situation and have Crop Reporting Service estimates of the size of the raisin-grape crop and can observe the current raisin market when they make their allocation decisions, they would not be likely to produce so many raisins that net returns would fall much below those in the crush market. Therefore, the no-volume-control simulations impose a restriction such that $PFD \geq K(PGCD)$, where K sets the minimum ratio of PFD to $PGCD$. If the initial simultaneous solution for grapes dried (QR) and crushed (QC) involving (3), (4), and (i) in block 2 results in $PFD < K(PGCD)$, (3) is eliminated. Equations (4) and (7), with DEL in place of QFR and DEL related to QR by a conversion ratio, are solved simultaneously for QR and QC subject to $QR + QC = Q$ and $PFD = K(PGCD)$. A value of $K = 7$ is assumed for the first no-volume-control run of the model. Because higher values of K might be appropriate if raisin growers viewed production as more risky without volume control, simulations with $K = 8$ and $K = 9$ are also run.

Yields, which are entered at historical annual values in all simulations, could be affected by the change in age distribution of vines resulting from changes in planting and removal rates without the reserve pool program; however, the effect is likely to be minor since yields are primarily determined by weather, diseases, and other natural factors. The quantity-to-fresh use, QF , which has averaged only about 10% of raisin grape production, could also vary in response to

changes in returns to dry and crush. However, because of costly changes in cultural practices required to switch from dry to fresh and the potential price effects in the fresh market, it is unlikely that this proportion would change significantly. Finally, without the volume control program, administrative costs of the marketing order would be reduced. A substantial portion of the grower assessment, which averages about 1% to 1.5% of grower costs, covers costs of providing market information to growers. To maintain a consistent environment, we assume the assessment for economic information would continue. The reduction in assessment resulting from elimination of reserve pool accounting would be such a small percentage of the unit cost of raisin production that no adjustment was made in the representative cost series.

No-Volume-Control Scenarios

With the above considerations, we have specified nine alternative scenarios that could plausibly reflect conditions without the reserve pool program.

Scenario 1: Planting and removal equations, (1) and (2), and the raisin-grape allocation equation (3) remain as in the historical model. The minimum ratio of the free tonnage price of raisins (PFD) to the grower price for raisin grapes crushed ($PGCD$) is set at 7 (i.e., $K = 7$). This restriction becomes binding in years of relatively large crops.

Scenario 2: Same as scenario 1 except planting response (including the disturbance) is reduced by 20% to reflect grower perceptions of increased risk without the marketing order.

Scenario 3: Same as scenario 1 except planting response is reduced 40% [roughly the findings of French and King (1988) for cling peaches].

Scenarios 1A, 2A, and 3A: Same as scenarios 1, 2, and 3 except (a) $K = 8$ to reflect a possible increase in perceived risk for raisin production without volume control and, therefore, a higher reservation price (effective in large-crop years) and (b) the coefficient of Q in equation (3) is reduced by .05 (from .6037 to .5887) to reflect an equivalent increase in risk perception for raisin production in years of small crops.⁶

Scenarios 1B, 2B, and 3B: Same as scenarios

⁶ Decreasing QR by .015 Q has roughly the same average effect on the allocation between drying and crushing as increasing K from seven to eight.

1, 2, and 3 except (a) $K = 9$ and (b) the coefficient of Q in equation (3) is reduced by .03 (from .6037 to .5737).

While no empirical observations are available to establish the range of scenarios A and B, the two variations provide a basis for evaluating both the general magnitude and direction of further changes in allocation response.

Simulation Results

Table 1 gives the time paths of three key indicators of the state of the industry with the volume control program in effect and as predicted without volume control, under three different planting response scenarios. For each scenario, bearing acres (BA) remain the same as under the volume control program for the first two years. However, without the reserve pool to absorb part of the raisin production, the grower price for raisin grapes allocated to drying ($PGRD$) and the representative measure of the average net return

to growers for raisin grapes dried and crushed (NGD) decline sharply.⁷ This leads to reduced plantings and slightly increased removals, which begin to affect bearing acres (and total production) in 1966. In 1967, a year of low yield and therefore no reserve pool allocations, the predicted grower price and net return per ton are higher under the no-volume-control scenarios than actual levels because of lower production from the reduction in bearing acres.

As time moves forward, the reduced returns of the first few years without the reserve pool lead to continued reductions in plantings and

⁷ $PGRD$ is a weighted average of the deflated free tonnage price and the grower export price. That is, $PGRD = [PFD(QP + QB) + (PGX/GNPD) \cdot QXT] \cdot C / (QP + QB + QXT)$, where C is the conversion ratio between metric tons of raisins (packed weight basis) and short tons of grapes used for raisins. Without the volume control program, $PGRD = PFD \cdot C$.

$NGD = [QR(PGRD - GCRD) + QC(PGCD - GCWD)] / (QR + QC)$, where $GCRD$ and $GCWD$ are representative unit costs of producing grapes for raisins. Since our cost estimates may not reflect the precise level of average industry experience, return measures are only indicators.

Table 1. Model Predictions of Three Indicators of the State of the Raisin Industry, 1964–85: Actual Values Under the Volume Control Program Compared with Values Under Three Levels of Planting Response without Controls

Year	Bearing Acres (BA), 1,000s				Grower Average Price for Raisin Grapes to Dry ($PGRD$)				Grower Average Net Return to Dry and Crush (NGD)			
	Actual	S_1	S_2	S_3	Actual ^a	S_1	S_2	S_3	Actual ^a	S_1	S_2	S_3
(1971–72 \$ per short ton)												
1964	252	252	252	252	73.5	63.1	63.1	63.1	7.4	-2.6	-2.6	-2.6
1965	254	254	254	254	66.0	37.2	37.2	37.2	-4.5	-25.7	-25.7	-25.7
1966	256	254	254	254	67.0	34.2	34.2	34.2	-10.9	-34.5	-34.5	-34.5
1967	255	252	251	250	81.2	96.3	97.0	97.7	0.9	4.2	4.5	4.8
1968	251	246	244	242	78.2	61.8	62.7	63.5	2.1	-8.5	-7.7	-7.0
1969	251	243	241	238	78.8	67.5	68.6	69.6	-2.7	-12.1	-11.2	-10.4
1970	246	235	233	231	71.7	76.9	77.8	78.7	-3.6	-3.0	-2.3	-1.5
1971	243	231	229	227	67.8	80.4	82.3	84.2	-6.5	-1.9	-1.0	-0.1
1972	241	228	226	224	116.8	148.8	151.2	153.9	16.2	25.3	26.0	26.7
1973	239	225	224	222	146.7	188.3	191.8	195.4	22.1	40.4	41.9	43.4
1974	241	227	225	223	121.7	122.7	125.9	129.2	-1.0	2.9	4.8	6.9
1975	241	226	224	222	104.7	86.6	89.5	92.7	-6.1	-14.7	-12.7	-10.4
1976	241	228	225	221	167.5	172.3	174.9	177.8	33.3	37.2	38.9	40.7
1977	243	232	227	223	130.7	136.9	140.3	143.9	17.1	20.7	23.1	25.7
1978	245	236	231	225	231.7	240.2	243.0	246.1	69.0	70.7	71.8	73.0
1979	246	238	233	227	150.9	150.8	156.1	161.8	21.9	22.1	25.9	30.0
1980	247	240	234	228	125.1	116.3	121.5	127.2	14.2	7.7	12.0	16.7
1981	253	245	239	232	150.9	166.4	174.2	183.0	5.9	14.7	19.5	24.8
1982	264	257	248	239	106.6	132.1	138.2	144.8	-23.7	-4.2	1.0	6.8
1983	277	272	260	248	113.8	61.1	66.0	77.7	-22.9	-66.2	-61.5	-51.4
1984	285	277	265	251	69.0	45.7	53.0	66.5	-71.3	-83.6	-78.1	-68.9
1985	284	274	261	248	68.4	39.9	47.3	59.8	-57.4	-76.2	-70.1	-60.5

^a If reserve pool sales of raisins in noncommercial channels are included in the grower average price computations with the unknown return arbitrarily valued at a deflated price of \$55 per short ton (approximately the price at which raisins were sold to packers under the export incentive program in effect in the early 1980s), the mean values of the grower return variables in the "actual" column are reduced as follows: $PGRD$ from 108.6 to 104.0 and NGD from 0.0 to -1.8. Deflated values may be expressed in 1985 dollars by multiplying by 2.35.

higher removals. Bearing acres fall about 15,000 below the historical level by 1975 for scenario 1 and slightly more under the reduced planting response assumed for scenarios 2 and 3. The differences in *BA* between actual and no-volume-control values narrow somewhat in later years under scenario 1, but widen under scenarios 2 and 3 which assume reduced planting response without the volume control program. Predicted production changes are proportional to bearing acres and variable yields, but under the no-control scenarios, relatively higher shares of raisin grape production are allocated to crushing in most years. By the decade of the 1970s, the reduced acreage and production under the no-volume-control scenarios bring the average grower return to or above the general level of returns with controls. However, returns without volume control again fall below the historical values in 1983–85.

If removal rates should increase without volume control, bearing acres would be reduced further in the early years, with some later modification resulting from feedback into the planting response. Because it seems likely that the overall effects would be contained within the general range of the three planting response scenarios, no specific scenarios involving removal variations are presented.

Table 2 presents the 1964–85 means and standard deviations of actual and predicted values for a larger group of variables and for the full set of alternative no-volume-control scenarios.⁸ Among the nine scenarios, note that for a given value of *K*, means of bearing acres and the output variables decrease with assumed reductions in planting response (*S2*, *S3*) and correspondingly the mean levels of prices and grower average net returns increase. As *K* increases, the allocation of grapes to dry declines as that to crush increases, leading to increases in grower and packer prices for raisins, decreases in the grower price for grapes crushed, but to increases in the combined indicator of grower average net return to dry and crush.

Table 2 suggests the following main effects

on prices, net returns, sales, and stocks of terminating the volume control program:

(a) The mean value of the average grower price for raisin grapes used for drying (*PGRD*) declines under scenario 1 but increases under all other scenarios. The standard deviation of *PGRD* increases about 30%.

(b) The mean price received by growers for raisin grapes crushed (*PGCD*) decreases slightly under most scenarios; the standard deviation increases only about 8%–10%.

(c) The mean value the deflated grower net return per ton of grapes, as measured by *NGD*, is reduced for all scenarios except where planting response is reduced by 40% (scenarios *S3*, *S3A*, *S3B*). The increase in standard deviation ranges from 16% for scenario 3B to 29% for scenario 1. Five years, 1965, 1966, 1983, 1984, and 1985 show losses much greater than ever experienced under the volume control program (see table 1).

(d) The mean value of the average FOB price received by packers (*PD*) increases slightly in scenario 1, but more in scenarios 2 and 3 under reduced planting response. The standard deviation of *PD* increases in the range of 17% to 24%.

(e) Most of the reduction in raisin production under the no-control scenarios was felt in exports rather than domestic shipments. Exports decreased in the range of 10% to 23%, depending on the scenario. The mean level of total carryover stocks decreased without volume control, but all these stocks would be held by packers since there would be no reserve pool (*CI* = 0). (Under volume control, mean level carryover stocks held by packers was $\overline{SP} = 29.4$, while that held by growers in the reserve pool was $\overline{CI} = 20.0$.)

(f) The packer net profit margin (*NPM*), computed by subtracting measures of representative unit packing cost and the cost of the raw product from the FOB packer prices in packaged and bulk form, was both lower and less variable under the volume control program.⁹ As may be observed from equations (8) and (9), the profit margin is affected by the ratios of raisin movement (Q^P , QB) to total raisin supply (SU). Under the marketing order, the total supply is larger relative to movement than without controls. In view of the potential market impact of this larger supply, packers set FOB prices rel-

⁸ In addition to the measures presented in table 2, approximate changes in economic surplus, as indicated by the areas under the demand curves facing processors, could be calculated. However, since over half of the raisins are sold in bulk for use in other food products and for export, interpretation of the surplus measures would be difficult. The value of such calculations is further limited because the model does not encompass the demand and supply systems for alternative crops grown on land shifted from raisin production. A complete examination of the demand, supply and marketing system for wine would also be needed.

⁹ $NPM = [NPMPCQP + NPMB(QB)] / (QP + QB)$, where $NMPM = PPD - PCPD - PFD$, $NPMB = PPD - PCBD - PFD$, and $PCPD$ and $PCBD$ are measures of packing cost per metric ton.

Table 2. Means (Standard Deviations) of Selected Variables Under Volume Control and Alternative No-Volume-Control Scenarios, 1964-85

Variable (s. t. = short ton, m. t. = metric ton)	Actual Data	No-Volume Control Scenarios								
		K = 7			K = 8					
		S1	S2	S3	S1A	S2A	S3A	S1B	S2B	S3B
Bearing acres (BA) (1,000 acres)	252 (13.6)	244 (15.8)	240 (13.6)	236 (12.3)	246 (15.7)	242 (13.4)	237 (11.8)	248 (15.9)	244 (13.3)	239 (11.5)
Grapes to dry (QR) (1,000 s. t.)	1,129 (317)	1,026 (327)	1,008 (316)	985 (298)	(1,013) (325)	995 (314)	972 (298)	998 (325)	981 (313)	959 (299)
Grapes to crush (QC) (1,000 s. t.)	740 (201)	774 (216)	755 (229)	741 (236)	805 (210)	785 (223)	768 (233)	836 (203)	815 (217)	795 (230)
Grower avg. price for grapes to dry (PGRD) (1971-72\$ per s. t.)	108.6 (42.7)	105.7 (56.0)	108.9 (56.6)	113.1 (56.8)	109.3 (54.5)	112.6 (55.2)	116.7 (55.5)	113.4 (53.8)	116.5 (54.2)	120.5 (54.5)
Grower price for grapes crushed (PGCD) (1971-72\$ per s. t.)	59.1 (19.2)	57.6 (21.2)	58.4 (21.0)	59.1 (21.0)	56.2 (21.1)	57.1 (20.9)	57.9 (20.9)	54.8 (20.8)	55.8 (20.7)	56.7 (20.8)
Avg. grower net return to dry and crush (NGD) (1971-\$ per s. t.)	0.0 (28.8)	-4.0 (37.2)	-1.7 (36.4)	1.2 (34.7)	-3.2 (36.2)	-0.9 (35.4)	1.9 (34.0)	-2.4 (35.4)	-0.1 (34.6)	2.6 (33.3)
Avg. FOB packer price for packaged and bulk NTS (PD)	896 (375)	907 (455)	933 (461)	967 (464)	933 (445)	960 (452)	993 (456)	963 (439)	988 (443)	1,020 (449)
Domestic shipments of packaged and bulk NTS (QD) (1,000 m. t.)	118.2 (21.4)	118.0 (25.7)	117.1 (25.3)	115.8 (24.4)	117.1 (25.7)	116.1 (25.3)	114.9 (24.5)	116.1 (25.6)	115.1 (25.3)	113.9 (24.6)
NTS exports (QXT) (1,000 m. t.)	45.6 (13.0)	39.4 (15.3)	38.5 (15.2)	37.2 (14.8)	38.3 (15.0)	37.3 (14.8)	36.1 (14.5)	37.0 (14.8)	36.0 (14.6)	34.9 (14.2)
Total carryover stocks (SP + CI) (1,000 m. t.)	49.9 (37.0)	38.6 (41.3)	36.6 (38.8)	34.1 (34.8)	37.8 (40.1)	35.8 (37.5)	33.4 (34.1)	36.5 (38.4)	34.7 (36.1)	32.6 (33.3)
Packer net margin for NTS (NPM) (1971-72\$ per m. t.)	45 (128)	103 (160)	111 (161)	122 (160)	109 (158)	118 (159)	129 (159)	117 (157)	126 (158)	136 (159)

Note: The means for PGRD, PGCD, PD, NPM, and NGD are simple averages of annual values. The quantity-weighted means are slightly lower but the relationship among actual and scenario values remains essentially unchanged. Also see footnote a, table 1.

actively lower for given raw product prices. The packer net margin was also more variable without the volume control program (the standard deviation increased about 25%) because of an increased variability in FOB packer prices relative to grower prices directly related to the increased variability of supply (*SU*).

Conclusions and Policy Implications

As with any econometric model, the empirical results are affected by the choice of equation forms, the specifications of participant expectations, and the estimation methods. The estimated model is an acceptable representation of the raisin industry in that it is based on generally accepted theoretical criteria and plausible behavioral assumptions, the parameter estimates are consistent with theoretical expectations, and the model tracks the historical system reasonably well. The most challenging aspect of the study is the modification of the model to reflect conditions that would have existed without the volume control program in effect, since such conditions are not observable historically. A set of scenarios was specified to reflect a range of possible no-control conditions. While the validity of these scenarios cannot be tested, their plausibility provides a basis for reaching tentative conclusions about the social costs and benefits of the reserve pool program using the performance criteria proposed by the USDA study team.

Criterion (1): above-normal profits. The measure of grower net return (*NGD*) averaged zero under the volume control program, suggesting that profits were not above normal. However, the representative cost series, while indicating changes in costs over time, may not fully reflect industry-wide averages. As table 1 shows, net returns varied considerably over time, and were affected by many factors besides the volume control program.

Criterion (2): price variability. Prices were less variable under the volume control program than for any of the no-control scenarios.

Criterion (3): disproportionate burdens. This analysis provides no information about possible

disproportionate burdens on particular classes of farmers.

Criterion (4): chronic surpluses. The average level of production was higher under volume control than for any of the no-control scenarios. This extra production can be viewed as "surplus" only if it results in returns persistently below competitive levels or continually requires low-use disposal of part of the crop. Grower returns under volume control, when most reserve pool quantities were either exported or carried over to be sold later, were above returns under most no-control scenarios. However, the higher average quantities carried from one year to the next ($SP + CI$) under volume control suggest a small increase in real marketing cost, which was not measured in the analysis. Significant diversions to noncommercial outlets (school lunches, needy persons, livestock feed, distilling) occurred in years of large crops, averaging about 4% of raisin production overall. These diversions could be regarded as surpluses, but perhaps not "chronic" in the sense intended by the USDA team.

Criterion (5): waste of resources. Resources used to produce raisins diverted to noncommercial use could be viewed as "wasted" because the returns generated are usually below cost. Most such diversions were in years of low or negative return as measured by *NGD* (table 2). The USDA study team argued that a marketing order may reasonably grant producers the right to withhold or divert quantities that would drive returns further below costs. Therefore, the diversions could be viewed as adjustments to unplanned short-run output variations.

Criterion (6): net revenue to growers. While the variable *NGD* is an imprecise measure of the level of grower net returns, relative comparisons indicate that mean returns were higher under the marketing order except for Scenarios 3, 3A, and 3B which assume substantial reductions in raisin supply response. Hence, it would be difficult to conclude that average grower net revenue was reduced because of the marketing order.

Criterion (7): consumer prices. Because the mean of the average FOB packer price (*PD*) is either about the same or above the historical mean for all no-control scenarios, the analysis sug-

gests that the marketing order program did not result in higher average prices to consumers, irrespective of the level of cost.

The mostly static theoretical studies of price stabilization have been generally critical of such programs. The dynamic empirical analysis presented here indicates that consumers were not made worse off in terms of average prices paid and quantities purchased (they may have been better off) and that the program reduced the variability of prices and returns. Grower average net returns may or may not have been increased, depending on the no-control scenario considered, but the marketing order clearly provided some protection in years of unexpectedly large production relative to demand. Overall, the results of the study suggest that the public interest may have been well served by the raisin volume control program, or at worst, there was no significant welfare loss.

[Received February 1990; final revision received August 1990.]

References

- Armbruster, Walter J., and Edward V. Jesse. "Fruit and Vegetable Orders." *Federal Marketing Programs in Agriculture: Issues and Options*, ed. W. J. Armbruster, D. R. Henderson, and R. D. Knutson, chap. 5. Danville IL: Interstate Printers and Publishers, 1983.
- French, Ben C. "Farm Price Estimation When There Is Bargaining: The Case of Processed Fruit and Vegetables." *West. J. Agr. Econ.* 12(1987):17-26.
- . "Fruit and Vegetable Marketing Orders: A Critique of the Issues and State of Analysis." *Amer. J. Agr. Econ.* 64(1982):916-23.
- . "Fruit and Vegetable Marketing Orders in the United States, 1937-1987: A Review." *Acta Horticulture* 223(1988):48-57.
- French, Ben C., and R. G. Bressler. "The Lemon Cycle." *J. Farm Econ.* 44(1962):1021-36.
- French, Ben C., and Gordon A. King. "Demand and Price-Markup Functions for Canned Cling Peaches and Fruit Cocktail." *West. J. Agr. Econ.* 11(1986):8-18.
- . *Dynamic Economic Relationships in the California Cling Peach Industry*. Giannini Foundation of Agr. Econ. Res. Rep. No. 338, University of California, 1988.
- French, Ben C., Gordon A. King, and Dwight D. Minami. "Planting and Removal Relationships for Perennial Crops: An Application to Cling Peaches." *Amer. J. Agr. Econ.* 67(1985):215-23.
- Heifner, Richard, Walter Armbruster, Edward Jesse, Glenn Nelson, and Carl Shaffer. *A Review of Federal Marketing Orders for Fruits, Vegetables, and Specialty Crops: Economic Efficiency and Welfare Implications*. Washington DC: U.S. Department of Agriculture, Agr. Mktg. Serv. Agr. Econ. Rep. No. 477, 1981.
- Jesse, Edward V. "Economic Efficiency and Marketing Orders" *Economic Efficiency in Agricultural and Food Marketing*, ed. Richard L. Kilmer and Walter J. Armbruster, pp. 217-28. Ames: Iowa State University Press, 1987.
- Kinney, William, Hoy Carman, Richard Green, and John O'Connell. *An Analysis of Economic Adjustments in the California-Arizona Lemon Industry*. Giannini Foundation of Agr. Econ. Res. Rep. No. 337, University of California, 1987.
- Lucas, Robert E., Jr. "Economic Policy Evaluation: A Critique" *The Phillips Curve and Labor Market*, ed. K. Bruner and A. H. Meltzer. Amsterdam: North-Holland Publishing Co., 1976.
- Minami, Dwight D., Ben C. French and Gordon A. King. *An Econometric Analysis of Market Control in the California Cling Peach Industry*. Giannini Foundation Monograph No. 39, Berkeley, University of California, 1979.
- Nuckton, Carole F., Ben C. French, and Gordon A. King. *An Econometric Analysis of the California Raisin Industry*. Giannini Foundation of Agr. Econ. Res. Rep. No. 339, University of California, 1988.
- Polopolus, Leo C., Hoy F. Carman, Edward V. Jesse, and James D. Shaffer. *Criteria for Evaluating Federal Marketing Orders: Fruits, Vegetables, Nuts and Specialty Crops*. Washington DC: National Technical Information Service, Identification Section, 1986.
- Thor, Peter K., and Edward V. Jesse. *Economic Effects of Terminating Federal Marketing Orders for California-Arizona Oranges*. Washington DC: U.S. Department of Agriculture, Econ. Res. Serv. Tech. Bull. No. 1664, 1981.
- U.S. General Accounting Office. *The Role of Marketing Orders in Establishing and Maintaining Orderly Marketing Conditions*. GAO/RCED-85-57, Washington DC, 1985.
- Wohlgenant, Michael J. *An Econometric Model of the U.S. Wine Industry*. Texas A&M University B-1507, 1985.

Are Farrowing Intentions Rational Forecasts?

David E. Runkle

This paper tests whether quarterly data on hog farmers' sow-farrowing intentions released in *Hogs and Pigs* are rational forecasts of actual farrowings. The empirical results show that neither one- nor two-quarter-ahead intentions are rational forecasts. Econometric methods and the implications of the empirical results for the rational expectations hypothesis are also discussed.

Key words: expectations, livestock forecasts, rationality tests.

Every three months the U.S. Department of Agriculture (USDA) releases *Hogs and Pigs*, a summary of the current livestock inventory and breeding intentions of the nation's hog farmers. Farmers, meat processors, and commodities traders examine this report because the news it contains affects their perceptions about future changes in hog prices. Most of the data in the survey reflect the current state of farmers' livestock herds. But the survey also reports the number of sows farmers intend to farrow in each of the next two quarters. Every issue of *Hogs and Pigs* releases the aggregate farrowing intentions of farmers in ten states.

Because hog farmers' farrowing intentions determine their current plans to buy or sell livestock, any discrepancy between their intentions and their subsequent actions can have a direct effect on their profits. Therefore, farmers should be expected to make the best predictions they can about their own future farrowing plans.

Farmers' farrowing intentions thus provide a natural test of the rational expectations hypothesis. Since the work of Muth thirty years ago (1961), some economists have based their models on the assumption that people have rational expectations about future economic events; that is, that people make the best forecasts they can, given what they know. Many authors have tried to determine whether people actually have rational expectations by testing whether the errors in forecasts of aggregate economic time series are systematic. But their critics point out that most people who answer these forecasting sur-

veys have nothing to lose if their forecasts are wrong. These critics argue that any test of the rational expectations hypothesis based on survey data should be dismissed.

But this criticism may appear to have no effect on the validity of testing the rationality of the farrowing intentions data from *Hogs and Pigs* because of the costs that farmers face if they make inaccurate predictions about their own behavior. Farmers may lose money if they make incorrect predictions about their future farrowing plans because it may be costly to change their plans.

But do farmers accurately report their own farrowing intentions? This paper shows that they do not. It shows that farmers' sow-farrowing intentions, as reported by the USDA, are not rational forecasts of farmers' actual sow-farrowing decisions. But the irrationality of these farrowing intentions does not necessarily refute the rational expectations hypothesis because farmers have few incentives to accurately report their intentions.

Other authors have long been concerned with the accuracy of government crop forecasts. However, the results in this paper are the first published tests of the rationality of government announcements of farmers' intentions for production of either livestock or crops. They extend previous work in agricultural economics on the accuracy of government crop forecasts.

Moore's 1917 analysis of government forecasts of cotton production was the first published work to examine the accuracy of government crop forecasts. Gunnelson, Dobson, and Pamperin examined the accuracy of USDA crop forecasts from 1929 to 1970. However, neither of these studies provided a formal statistical test

David E. Runkle is a senior economist in the Research Department of the Federal Reserve Bank of Minneapolis and an adjunct associate professor of finance at the University of Minnesota.

of the unbiasedness of government crop forecasts. Sumner and Mueller come closest to the tests reported here. In a footnote, they discussed statistical analysis of the accuracy of crop forecasts. But it is not clear whether they performed formal tests of unbiasedness. No tests of the efficiency of crop forecasts or intentions announcements appear to have been published.

How Can the Rationality of Farmers' Expectations Be Tested?

Suppose that a farmer has determined how many sows he will farrow k quarters from now. Denote the current quarter time t and the particular farmer, farmer i . Farmer i makes a prediction at time t about the number of sows he will farrow at time $t + k$. His prediction will be called ${}_i f_{i,t+k}^p$, and the actual number of sows he farrows in time $t + k$ will be called $f_{i,t+k}$.

If the farmer's prediction is rational, it is equal to the mathematical expectation of $f_{i,t+k}$, conditioned on what he knew at time t ; that is, ${}_i f_{i,t+k}^p = E(f_{i,t+k} | I_{i,t})$, where $I_{i,t}$ is the information available to farmer i at time t , and E is the mathematical expectations operator. Forecast rationality also implies that the forecast error from the farmer's prediction cannot itself be predicted using information known when the forecast was made.

If they were available, data on the farrowing intentions and actual farrowings of a particular farmer could be used to test whether his intentions were rational forecasts of his actual decisions by running the following regression:

$$(1) \quad f_{i,t+k} = \alpha_0 + \alpha_1 {}_i f_{i,t+k}^p + \beta' X_{i,t} + \epsilon_{i,k}^i$$

where $X_{i,t}$ is any vector of variables in farmer i 's information set at time t , and $\epsilon_{i,k}^i$ is the error term.

Unbiasedness requires that, in a regression without $X_{i,t}$ variables, the coefficients in equation (1) may be restricted to $\alpha_0 = 0$ and $\alpha_1 = 1$. Efficiency imposes the additional requirement that any variable known at time t or before be orthogonal to $\epsilon_{i,k}^i$; that is, the vector $\beta = 0$ for any vector $X_{i,t} \in I_{i,t}$. A farmer's farrowing intentions are rational only if they are both unbiased and efficient.

There are two reasons why tests of rationality cannot be conducted in this simple manner. First, the data in *Hogs and Pigs* do not reveal the intentions or actions of individual farmers; instead, they provide only aggregate data. Sec-

ond, the failure of certain standard statistical assumptions implies that inference based on ordinary least squares regression (OLS) will not be valid, because the OLS estimate of the covariance matrix will be incorrect.

The discussion of equation (1) presented the restrictions implied by forecast rationality for an individual farmer. Suppose N farmers are sampled in each period, and all of those farmers make rational predictions about their own sow-farrowing decisions. In this case, aggregate farrowing intentions should be a rational forecast of actual aggregate farrowing. If aggregate farrowing intentions at time t are denoted as ${}_t F_{t+k}^p = \sum_{i=1}^N {}_i f_{i,t+k}^p$ and aggregate actual farrowings at time $t + k$ are denoted as $F_{t+k} = \sum_{i=1}^N f_{i,t+k}$, then, because of linearity, rationality implies the same coefficient restrictions in the equation

$$(2) \quad F_{t+k} = \alpha_0 + \alpha_1 {}_t F_{t+k}^p + \beta' X_t + \epsilon_{t,k}$$

as it does in equation (1).¹ Therefore, if each individual farmer makes a rational forecast, the aggregate forecast will be rational, as well.

Although individual rationality implies aggregate rationality, the converse is not true. If one farmer always underpredicts his farrowing by fifty sows while another always overpredicts his farrowing by fifty sows, the sum of their forecasts would be unbiased even though neither farmer makes an unbiased forecast. This example illustrates why aggregate data can disprove individual rationality but cannot prove it.

The problems with standard OLS statistical assumptions are somewhat less transparent than the issues of aggregation. Inference based on OLS regressions assumes that the errors in the regression equation are independent over time and have the same variance in each period. Each of these assumptions may be invalid in estimating equation (2). If either assumption is violated, then the OLS covariance matrix will not be a consistent estimate of the true uncertainty in the estimated parameters in equation (2), even though the OLS parameters themselves are consistent estimates of the true parameters.

The first problem with using OLS to estimate equation (2) is that the errors in that equation may be serially correlated. If farmers predict farrowings two periods ahead, then $k = 2$ in

¹ The only difference between the two equations is the interpretation of X_t . In equation (2), X_t can be either data that was publicly known at time t or the sum across farmers or particular pieces of private information that each farmer had. For example, since each farmer would know how many sows he had farrowed in time t , the linearity of equation (1) implies that the aggregate number of sows farrowed in t can be included in X_t .

equation (2). In this case, the farmer's forecast errors can be serially correlated even if their forecasts are rational. Figure 1 demonstrates how this serial correlation could arise and shows how to determine the order of the serial correlation.

Suppose that at time t , a farmer predicts his sow-farrowing intentions for time $t + 2$, $f_{i,t+2}^p$. At time $t + 2$ when his actual farrowings, $f_{i,t+2}$, are known, his two-period-ahead forecast error, $\epsilon_{i,2}^i = f_{i,t+2} - f_{i,t+2}^p$, is also revealed. If the farmer's farrowing intentions announcement is a rational forecast of his actual farrowings, his forecast error should be uncorrelated with anything he knew when he announced his intentions.

For example, figure 1 shows that the error the farmer made at time $t - 2$ in predicting his sow farrowing at time t , $\epsilon_{i,t-2}^i = f_{i,t} - f_{i,t-2}^p$, is known when he announces his intentions at time t for time $t + 2$. Therefore, $\epsilon_{i,2}^i$ should not be correlated with $\epsilon_{i,t-2}^i$ if his farrowing intentions announcement is a rational forecast, because $\epsilon_{i,t-2}^i$ is in the information set $I_{i,t}$.

Figure 1 also shows that the farmer does not know the error he made at time $t - 1$ in predicting his farrowings in time $t + 1$. That error,

$\epsilon_{i,t-1,2}^i = f_{i,t+1} - f_{i,t-1}^p$, becomes known only at time $t + 1$ when actual farrowings for that period, $f_{i,t+1}$ are known. Therefore, even if the farmer's intentions announcement were a rational forecast of his actual farrowings, the farmer's forecast error would be serially correlated because $\epsilon_{i,t-1,2}^i$ is not known when $f_{i,t+1}$ is announced. That is, $\epsilon_{i,t-1,2}^i$ is not in $I_{i,t}$.

Thus, if the farmer's intentions announcement is a rational forecast, $\epsilon_{i,2}^i$ can be correlated with $\epsilon_{i,t-1,2}^i$, but it cannot be correlated with $\epsilon_{i,t-2,2}^i$. This correlation implies that the errors in equation (1) would be moving-average errors of order 1 [MA(1)]. Hansen and Hodrick first demonstrated that this kind of forecast-error correlation would arise in multiperiod forecasts, even if the predictions were rational. The aggregate data in equation (2) should have exactly the same error structure as the individual data if all farrowing intentions are rational forecasts.

Although the errors in equation (2) would be serially correlated if farmers were making two-period-ahead forecasts ($k = 2$), using a generalized least-squares (GLS) transformation to eliminate the serial correlation in that equation would be invalid because the principal regressor

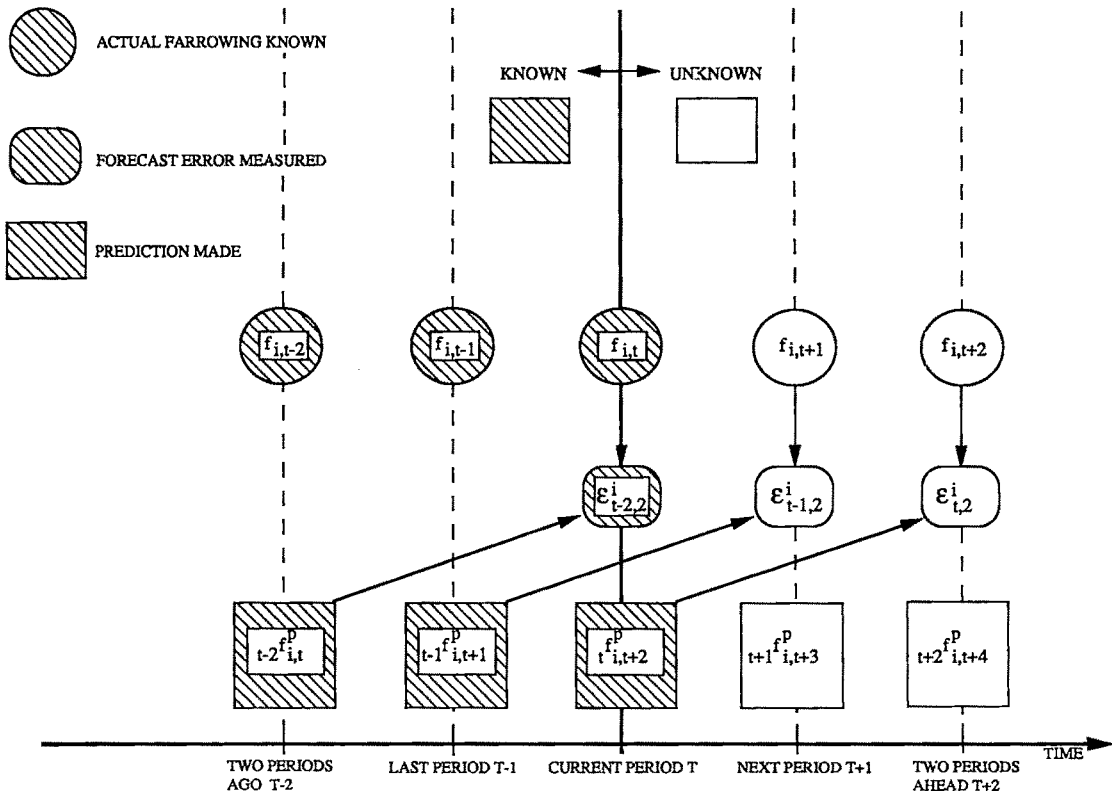


Figure 1. What a farmer knows when predicting his sow farrowings two periods later

(F_{t+2}^p) is predetermined, rather than exogenous. In fact, as Hansen and Sargent show, conventional GLS estimates of the parameters in (2) would be inconsistent because the transformed errors would be correlated with the transformed regressors.² This problem would be even more serious than the misestimation of the covariance matrix that would result from using an OLS regression.

If farmers were making only one-period-ahead forecasts ($k = 1$), moving-average errors would not be present in equation (2). Nevertheless, inference using OLS could still be invalid if the conditional variance of the error were correlated with the regressors. This problem, which White calls conditional heteroskedasticity, often arises in time-series data. If conditional heteroskedasticity exists in the data, then the OLS estimate of the covariance matrix will be incorrect even if there are no moving-average errors.

The problems in inference created both by moving-average errors and conditional heteroskedasticity can both be solved by using Hansen's generalized-method-of-moments (GMM) estimator to perform tests of forecast rationality. Hansen's estimator uses a set of optimally weighted orthogonality conditions to estimate parameters efficiently for most standard econometric models. For equation (2), the orthogonality conditions are that the estimated error is orthogonal to each of the regressors, that is,

$$E((F_{t+k} - (\alpha_0 + \alpha_1 F_{t+k}^p + \beta' X_t))' Z_t) = 0, Z_t \in \{1, F_{t+k}^p, X_t\}.$$

Hansen shows that any GMM estimator takes the following form:

$$(3) \quad \min_{\gamma} g(\gamma)' W g(\gamma),$$

where γ is a vector of coefficients, $g(\gamma)$ is the vector of sample averages of the set of orthogonality conditions specified in a given statistical model, and W is a weighting matrix. In equation

(2), the vector γ is the set of coefficients α_0 , α_1 , and β .

Hansen also shows that the optimal weighting matrix W for a GMM estimator is simply the inverse of the covariance matrix of the orthogonality conditions for that particular equation. If the errors for a particular equation are $MA(p)$, then the estimate of the covariance matrix of the orthogonality conditions is

$$(4) \quad \hat{\Omega} = \frac{1}{T} \sum_{j=-p}^p \sum_{t=1}^T Z_t' \hat{\epsilon}_{t,k} \hat{\epsilon}_{t-j,k}' Z_{t-j},$$

where $\hat{\epsilon}_{t,k} = ((F_{t+k} - (\hat{\alpha}_0 + \hat{\alpha}_1 F_{t+k}^p + \hat{\beta}' X_t))' Z_t)$, and consistent estimates of the parameters are used to form the residuals used in constructing the weighting matrix.

If $\hat{W} = \hat{\Omega}^{-1}$, then the GMM estimator is efficient within the class of estimators that uses only conditional moment restrictions, as was shown by Chamberlain. Because of the moving-average error structure, this estimate is not constrained to be positive definite. Newey and West discuss consistent estimates that are guaranteed to be positive definite. The Newey-West estimator is used in the GMM results presented in this paper.

For this analysis, the GMM parameter estimates are exactly the same as the OLS parameter estimates because equation (2) is exactly identified: the number of instruments and the number of regressors is the same. But the GMM standard errors will be correct in the presence of moving-average errors and conditional heteroskedasticity, while the OLS standard errors will be incorrect. Almost all empirical tests in this paper are therefore based on GMM estimation.

Data and Empirical Results

For almost thirty years, *Hogs and Pigs* has published quarterly aggregate data on sow-farrowing intentions for farmers in ten states.³ These data are gathered through interviews conducted by the USDA. Starting on the first day of March, June, September, and December, USDA interviewers ask farmers how many sows they intend

² Consider the case of testing the unbiasedness of equation (2) when $k = 2$. The equation of interest would be $F_{t+2} = \alpha_0 + \alpha_1 F_{t+2}^p + \epsilon_{t,2}$. Because the error term would have an $MA(1)$ structure under the hypothesis of unbiasedness, assume that that error structure can be represented as $\epsilon_{t,2} = \eta_{t+2} - \gamma \eta_{t+1}$. GLS would require that we filter equation (2) by $(1 - \gamma L)^{-1}$ to get rid of the serial correlation in $\epsilon_{t,2}$. However, this would cause the filtered equation to violate the OLS orthogonality conditions. Note that the transformed regressor would be $F_{t+2}^p = (1 - \gamma L)^{-1} F_{t+2}^p = (F_{t+2}^p + \gamma_{t-1} F_{t+1}^p + \gamma_{t-2}^2 F_{t+2}^p + \dots)$, while the transformed error term would be $\tilde{\epsilon}_{t,2} = (1 - \gamma L)^{-1} \epsilon_{t,2} = (\epsilon_{t,2} + \gamma \epsilon_{t-1,2} + \gamma^2 \epsilon_{t-2,2} + \dots)$. However, we know that $F_{t+2}^p \perp \epsilon_{t,2}$ if, and only if, $j \leq k$. But since the product $F_{t+2}^p \epsilon_{t,2}$ would contain terms like $F_{t+2}^p \epsilon_{t,2}$ where $j > k$, $F_{t+2}^p \not\perp \epsilon_{t,2}$.

³ Until 1968, the title of the report was *Pig Crop Report*. Since 1940, this report has also contained biannual intentions data for farmers in the entire country. The USDA increased the number of quarterly states to fourteen in 1973 but continued to publish data for the original ten states. In 1982 the USDA changed its ten-state sample. Later in this section, the effect of these changes on the tests of rationality is discussed.

to farrow in the next three months, how many sows they intend to farrow in the three months after that, and how many sows they farrowed in the preceding three months.⁴ The survey methods for *Hogs and Pigs* are described in USDA (1977, 1979).

The empirical tests in this paper use aggregate published data on quarterly farrowings in ten states from the first quarter of 1961 until the first quarter of 1989.⁵ The goal of this paper is to test whether the first and second farrowings intentions announcements are rational forecasts of actual farrowings. To see how these tests were conducted, consider the following example. In September 1960 and December 1960, farmers were asked how many sows they planned to farrow between 1 December 1960 and 28 February 1961. In March 1961, farmers were asked how many sows they actually farrowed in the previous three months. The analysis here compares the September and December intentions data to the actual farrowings announced the following March.

Two potential problems with the data need to be addressed before the empirical findings are reported. First, inference using the entire sample could be invalid because the ten-state sample changed in 1982. Second, inference using the initial announcements of actual farrowings could be invalid if those announcements were biased.

To solve the first problem, the observations which would have resulted in comparing intentions data for one set of ten states with actual farrowings for the other set were dropped. Even if the sample of states is changed, as long as the same farmers provide both the intentions data and the actual farrowings data, the tests of rationality are valid.⁶

The second problem was somewhat more difficult to address. If the initial announcements of actual farrowings were revised in a predictable way, then the data used here could result in incorrect inference. One way to determine whether such revisions could affect statistical inference is to check whether the initial announcements of sow-farrowing intentions are rational forecasts of the revised data. Such a test was developed by Mankiw, Runkle, and Shapiro. If the initial

announcement is a rational forecast of the revised data, then using the initial announcement will cause no problems in inference. In fact, GMM estimates show that the initial announcement is a rational forecast for any of the next three quarterly revisions. Therefore, data on the initial announcement of sow farrowings in the past quarter are used as the dependent variable.⁷

Because a prediction must be both unbiased and efficient to be rational, separate tests for unbiasedness and efficiency were conducted. Table 1 shows the results of the unbiasedness tests. The first column of that table shows that the USDA series on aggregate intentions of sow farrowings during the next quarter is an unbiased predictor of actual aggregate farrowings in ten states. Since the hypothesis of rationality imposes two restrictions, if the intentions data were unbiased, the test statistic reported in column 1 should be asymptotically distributed as a chi-square with 2 degrees of freedom random variable. Since its value is only 4.228, the hypothesis of unbiasedness cannot be rejected.

Note that the results in column 1 were based on a GMM estimate of the covariance matrix of the parameters. Column 2 shows that, if the OLS estimate of the covariance matrix were incorrectly used, the hypothesis of unbiasedness would have been incorrectly rejected. This result illustrates the importance of using GMM estimates of the covariance matrix in time-series models. All additional results in this paper come from GMM estimation.

Column 3 shows the results of a test of whether the USDA series on aggregate intentions of sow farrowings in the period three to six months following the survey is an unbiased predictor of actual farrowings during that period. The results in this column show that those intentions are biased predictors and are therefore not rational forecasts of actual farrowings. The chi-square test statistic shows that the hypothesis of unbiasedness can be rejected at the 1% confidence level. This rejection is strong evidence that the two-quarter-ahead announcement of sow-farrowing intentions is not a rational forecast of actual farrowings. In fact, if two-quarter-ahead farrowings were predicted using the regression reported in column 3 instead of the two-quarter-ahead farrowings intentions announcement, the

⁴ For example, in December 1989, farmers were asked "Of the sows and gilts [you are holding in inventory] how many are expected to farrow between March and May 1990," to determine their two-quarter-ahead farrowing intentions.

⁵ Because of the timing of the survey, the first quarter of 1961 is really December 1960 through February 1961.

⁶ As with all tests of rationality using aggregate survey data, the effects of sampling error are ignored.

⁷ The government announcements of actual farrowings are adjusted using data other than the survey. The rationality tests for the initial announcement are not reported here because they are not germane to determining whether farmers farrowing intentions, rather than government data revisions, are rational.

Table 1. Tests for Unbiasedness of One- and Two-Quarter-Ahead Aggregate Sow-Farrowing Intentions

	$F_{t+k} = \alpha_0 + \alpha_1 F_{t+k}^p + \epsilon_{t,k}$		
	(1) One Quarter ($k = 1$)	(2) One Quarter ($k = 1$)	(3) Two Quarters ($k = 2$)
Forecast horizon			
α_0	111.926 (58.026) ^a	111.926 (47.312)	228.628 (70.644)
α_1	0.947 (0.026)	0.947 (0.020)	0.901 (0.031)
Estimation method	GMM	OLS	GMM
Number of moving-average terms	0	0	1
χ^2 statistic for $H_0 : \alpha_0 = 0, \alpha_1 = 1$	4.228	8.581	10.518
Significance level	(0.121)	(0.014)	(0.005)
Number of observations	111	111	109

^a Standard errors are in parentheses under coefficients.

mean-square forecast error would be reduced by over 10%.

Because the two-quarter-ahead intentions announcements are biased, they are not rational forecasts; therefore, there is no point in testing whether they are efficient forecasts. But since the one-quarter-ahead intentions announcements appear to be unbiased, further tests are warranted to determine whether they are efficient forecasts.

One piece of information that each farmer knows when announcing his one-quarter-ahead intentions is his own two-quarter-ahead intentions from the previous period. In fact, his current one-quarter-ahead intentions are a revision of his previous intentions. If the current announcement of intentions is an efficient forecast, adding the previous intentions as an additional explanatory variable should not improve the accuracy of the forecast.

Table 2 shows a test of this hypothesis. Since β_1 is highly significant, lagged two-period-ahead intentions provide information useful in forecasting that is not included in the one-period-ahead intentions data. The chi-square test shows that the hypothesis of forecast efficiency can be rejected at the 1% level. Therefore, the one-quarter-ahead intentions report is not a rational forecast of actual farrowings because it is not an efficient forecast. If one-quarter-ahead farrowings were predicted using the regression reported in table 2 instead of the one-quarter-ahead farrowings intentions announcement, the mean-square forecast error would be reduced by over 11%.

What Does the Rejection of Forecast Rationality Mean?

These results show that neither the one-quarter-ahead nor the two-quarter-ahead farrowing-intentions announcement in *Hogs and Pigs* is a rational forecast of actual farrowings. But what does this mean? It could mean that farmers are not good at predicting their own actions, but this is probably not the most likely explanation. A farmer would lose money if he did not accurately predict his own future actions, but there is no cost to him if he does not accurately report his intentions.

Although it may be somewhat surprising that farmers announce irrational forecasts of their own future actions, it would be considerably more

Table 2. Test for Efficiency of One-Quarter-Ahead Aggregate Sow-Farrowing Intentions

$F_{t+1} = \alpha_0 + \alpha_1 F_{t+1}^p + \beta_1 F_{t-1}^p + \epsilon_{t,1}$	
α_0	113.647 (60.268) ^a
α_1	1.232 (0.119)
α_2	-0.286 (0.115)
Estimation method	GMM
Number of moving-average terms	0
χ^2 statistic for $H_0 : \alpha_0 = 0, \alpha_1 = 1, \beta_1 = 0$	10.813
Significance level	(0.004)
Number of observations	109

^a Standard errors are in parentheses under coefficients.

surprising if livestock market analysts were to announce irrational forecasts of farmers' actions. Because the market analysts, unlike farmers, are paid for the accuracy of the forecasts they report, they have a strong economic incentive to report accurately. In fact, Colling and Irwin recently found that the averages of market analysts' expectations about the announcements in *Hogs and Pigs* on reported changes in breeding- and market-hog inventories are rational forecasts. Thus, the different outcomes of these two rationality tests may stem from the different economic incentives that farmers and market analysts have to report their expectations accurately.

These differences in the rationality of reported expectations are paralleled in the economic literature on testing the rationality of inflation forecasts. Carlson and Mullineaux, among others, found that the inflation forecasts in the Livingston survey are not rational forecasts of actual inflation. In contrast, Keane and Runkle found that the price forecasts made by professional forecasters in the American Statistical Association-National Bureau of Economic Research Survey are rational forecasts. They suggest that the nonprofessional forecasters in the Livingston survey may have few economic incentives to accurately report their expectations, and thus, the irrationality of those reported expectations is not surprising.

Because farmers appear to have few economic incentives to report their farrowing intentions accurately, the fact that those intentions are not rational forecasts does not necessarily disprove the rational expectations hypothesis. Nevertheless, this paper has shown that forecasts of actual sow farrowings can be significantly improved by discarding the assumption that the aggregate sow-farrowings intentions announcement is unbiased and efficient.

[Received October 1989; final revision received September 1990.]

References

Carlson, John A. "A Study of Price Forecasts." *Ann. Econ. and Soc. Measure*. 6(1977):27-56.

- Colling, P. L. and S. H. Irwin. "The Reaction of Live Hog Futures Prices to USDA *Hogs and Pigs Reports*." *Amer. J. Agr. Econ.* 72(1990):84-94.
- Chamberlain, G. "Panel Data." *Handbook of Econometrics*, Vol. 1, ed. Z. Griliches and M. D. Intriligator, pp. 1247-1313. Amsterdam, New York: Elsevier Science Publishers, 1984.
- Gunnels, G., W. D. Dobson, and S. Pamperin. "Analysis of the Accuracy of USDA Crop Forecasts." *Amer. J. Agr. Econ.* 54(1972):639-45.
- Hansen, L. P. "Large Sample Properties of Generalized Method of Moments Estimators." *Econometrica* 5(1982):1029-54.
- Hansen, L. P., and R. J. Hodrick. "Forward Exchange Rates As Optimal Predictors of Future Spot Rates: An Econometric Analysis." *J. Polit. Econ.* 88(1980):829-53.
- Hansen, L. P., and T. J. Sargent. "Instrumental Variables Procedures for Estimating Linear Rational Expectations Models." *J. Monetary Econ.* 9(1982):263-96.
- Keane, M., and D. Runkle. "Testing the Rationality of Price Forecasts: New Evidence from Panel Data." *Amer. Econ. Rev.* 80(1990):714-35.
- Mankiw, N. G., D. E. Runkle, and M. D. Shapiro. "Are Preliminary Announcements of the Money Stock Rational Forecasts?" *J. Monetary Econ.* 14(1984):15-27.
- Moore, H. L. *Forecasting the Yield and Price of Cotton*. New York: Macmillan Co., 1917.
- Mullineaux, D. J. "On Testing for Rationality: Another Look at the Livingston Price Expectations Data." *J. Polit. Econ.* 82(1978):329-36.
- Muth, J. F. "Rational Expectations and the Theory of Price Movements." *Econometrica* 29(1961):315-35.
- Newey, W. K., and K. West. "A Simple, Positive Semidefinite, Heteroskedasticity and Autocovariance Consistent Covariance Matrix." *Econometrica* 55(1987):703-8.
- Sumner, D. A., and R. A. E. Mueller. "Are Harvest Forecasts News? USDA Announcements and Futures Market Reactions." *Amer. J. Agr. Econ.* 71(1989):1-8.
- U.S. Department of Agriculture, Economics, Statistics, and Cooperatives Service. *Hog and Pig Reports: A Handbook on Surveying and Estimating Procedures*. ESCS-66, Washington DC, 1979.
- U.S. Department of Agriculture, Statistical Reporting Service. *Scope and Methods of the Statistical Reporting Service*. SRS Misc. Pub. No. 1308, Washington DC, 1977.
- White, H. "A Heteroskedasticity-Consistent Covariance-Matrix Estimator and a Direct Test for Heteroskedasticity." *Econometrica* 48(1980):817-38.

The Welfare Analytics of Transaction Costs, Externalities, and Institutional Choice

Ronald C. Griffin

A natural result of the economist's participation in externality literature has been a strong emphasis upon market and price-guided policies. Available theoretical foundations are insufficient for supporting this ideology. Absent is a conceptual framework which integrates institutional options and their attendant resource costs with what we know of resource and technology constraints. This paper addresses this omission by incorporating an explicit role for transaction costs into traditional welfare diagrammatics.

Key words: externality, institutions, property rights, resource policy, transaction costs, welfare economics.

Externality and market failure literature of recent decades has been responsible for a substantial amount of discussion, debate, and insight regarding the social relations which bind and, indeed, define mankind. These developments have been spurred in no small part by the policy relevance of externality-like issues. In addition, because economists have believed that they offer valuable input for such issues, market devices for "correcting" externalities have always been emphasized in the literature, often to the exclusion of other social/legal relationships. "When the only tool you have is a hammer, everything looks like a nail." Throughout this literature, concern for the relative abilities of the market and government policy has usually been explicit.

One of the primary themes of the externality literature and its variants should be that a wide array of institutions is available for correcting externalities. While certain camps might focus on narrowly defined subsets of these internalizing institutions, common sense suggests that society examine externality situations on a case-by-case basis and choose the best structure available for accommodating the externality

interdependence. The importance of transaction costs becomes apparent in the pursuit of this objective.

A major tenet of the more recent literature has been the integral involvement of transaction costs within organizational processes which seek to have externality creators account for their impacts upon recipients. These processes include both market and nonmarket forms of internalization, and all such alternatives incur transaction costs to some degree. Literature on this subject observes that markets for accommodating externality problems will naturally arise if transaction costs are sufficiently low (Demsetz 1964, 1967). Moreover, because it is rarely possible to know, a priori, whether market-incurred transaction costs are exceeded by the transaction costs of each policy (nonmarket) alternative, the desirability of policy is usually founded upon subjective biases regarding relative levels of transaction costs (Dahlman).

While transaction costs are usually viewed as affecting the economic efficiency of alternative internalizing institutions, it is not typically acknowledged that transaction costs are also relevant in defining economic efficiency. The prevalence of transaction costs implies that concepts such as Pareto optimality are not immune to a consideration of information and its costs. Even when this fact is suggested, means for operationalizing it are rarely provided. By properly applying some traditional "welfare ana-

Ronald C. Griffin is an associate professor, Department of Agricultural Economics, Texas A&M University.

Texas Agricultural Experiment Station Paper No. TA-25794.

The author expresses appreciation to Andy Dragan, Wes Peterson, John Stoll, the *A/AE* reviewers and, especially, Dan Bromley for comments and suggestions.

lytics," an insightful background can be obtained for institutional choice in the presence of transaction costs. The purpose of this paper is to expand neoclassical tools of welfare analysis (summarized by Bator) to accommodate a role for transaction costs.

Definitions and Background

To facilitate the exposition it is useful to establish a few fundamental definitions. It is particularly important to elucidate the concepts of externality, property rights, and transaction costs so that a common basis for the forthcoming discussion can be obtained. With regard to externality, the following definition is adopted and progressively justified in this paper. An externality is an interdependence among people. Externality, as it is defined here, does not imply Pareto relevancy. That is, there is no presumption of social ill in this definition. This definition is not mainstream, for the literature more commonly defines externality as an interdependence among economic actors for which a market or some other compensatory device does not exist (Baumol and Oates, Bromley 1986, Heller and Starrett). This second definition offers instructional utility, which hopefully explains its predominance, but it has the potential to perform great damage in application. While economists have a predilection for emphasizing market absence as a necessary condition for externality, an institutionally unbiased definition allows one to entertain nonmarket allocation processes as potential externality remedies. Casual reflection will isolate numerous social modes of externality resolution. For the most part, these institutions are so broad and permeating that they escape exhaustive classification efforts and, often, even our abilities to identify particular mechanisms.¹ One purpose of this paper is

to assign a greater standing to the diverse array of nonmarket externality-addressing institutions. By directly incorporating transaction costs into the definition of Pareto optimality, the central objective of this paper is to build upon the idea that a Pareto-relevant externality is an externality that is being addressed by the wrong institution rather than not being addressed by any institution (especially markets) at all.

In the course of studying the entirety of institutions relevant to externality interdependencies, insight can be achieved by carefully emphasizing markets. The concept of property rights is crucial here, and, with Macpherson's warning in mind, "the meaning of property is not constant" (p. 1), it is useful to preface forthcoming discussion with a brief sketch of the elements composing property rights. Following Calabresi and Melamed's classic paper, property rights can take the form of property rules, liability rules, or inalienable entitlements.² Protection of one's interest in an interdependence by a property rule infers that one's desires in this relationship are ascendant to the desires of other involved parties unless the protected individual chooses to reassign this control. Similarly, protection via a liability rule in an interdependence provides for compensation should other individuals choose to assume control over the relationship. An inalienable entitlement declares that one's interests in a relationship cannot be legally reassigned or taken under any circumstances even when each party to the interdependence is willing.

Property rules are fundamental to the operation of markets which assign value prior to an interdependence, and liability rules necessitate some form of judicial system to settle value after an interdependence has occurred. Generally speaking, property rules are the foundation of market activity, and liability rules are the basis for litigation. However, sufficiently established

¹A few examples of common interdependence-internalizing institutions may be in order.

Envy is a term representing a specific class of consumption interdependencies among individuals. Envy, therefore, qualifies as an externality given either the above definition or the more traditional "interdependence without a market" definition. Institutions that address this type of interdependence do, however, exist. Society seeks to teach or convince individuals that envy is undesirable, even irrational. In this case the intent of the control institution is to repress the interdependence. Social teachings (e.g., the ninth and tenth Commandments) serve to allocate the costs of envy to the envious themselves. This is conceivably an efficient solution, because the envious may be in the best position to solve the problem—by adapting their utility functions to be less responsive to the interdependence.

A child is dependent on his or her parents for the provision of a positive environment in which to live, learn, and mature. Yet, no

market is available for the interdependence. Instead, parents possess a felt or declared responsibility to bear the burden of the interdependence. The source of these feelings may be partly founded on instinct, but this is assuredly reinforced by the parents' childhood participation in family relationships in addition to the important role of social teachings. It is notable that changes in tastes or relative scarcities may have recently led to dissatisfaction with the ability of family/community/church institutions to internalize the child-parent interdependence. As a result, additional institutions are being tried. Public awareness and opinion regarding child abuse has increased. Child abuse is being more broadly defined. Welfare agencies now act in an oversight capacity, monitoring the performance of parents previously suspected to be remiss in their parental duties. The "rights" of children are being refined to the point that they are better able to serve as the basis for court action brought upon parents by their children or by others.

²This classification is essential to a great deal of literature. See, for example, Bromley (1978), Buchanan and Faith, and Frech.

liability rules also may provide a basis for market transactions, and the ongoing refinement of property rules often occurs in the courthouse. Inalienable entitlements act as constraints on market activity, presumably to advance higher-order social objectives.

In forthcoming discussion, the emergence of markets to address an externality interdependence will be equivalent to the establishment of a property rule. In the simplistic view of this process, society must decide which person or group of persons should be vested with a newly created property rule. The more complex (and more realistic) perspective is that property rules do not emerge into an institutional vacuum but must coexist in an evolved system of, among other things, detailed property rules. Therefore, new property rules are best characterized as additional circumscriptions or alterations within social institutions. The simpler of these two interpretations is not too misleading for the more immediate purposes of this paper.

Transaction costs include "the costs of obtaining information, establishing one's bargaining position, bargaining and arriving at a group decision, and enforcing the decision made" (Randall 1972, p. 176n). Dahlman separates transaction costs into (a) search and information costs, (b) bargaining and decision costs, and (c) policing and enforcement costs and then states that all of these costs "represent resource losses due to lack of information" (Dahlman, p. 148). Moreover, the amount and distribution of transaction costs will differ between property rules and liability rules as well as between situations in which an initial entitlement (either a property or liability rule) is held by one party or another (Bromley 1978).

The Welfare Analytics of Property Rules

It is convenient to introduce the elements of this discussion in a progressive fashion beginning with the establishment of a general externality. Let us temporally adopt, for the sake of argument, the instructional definition: Externalities are interdependencies among economic actors which, for the purpose of the immediate discussion, are not accompanied by markets. A specific act of internalization merely causes the relation to enter into the decision calculus of causal economic agents who previously ignored it in full or in part. The interdependence may involve producers, consumers, or both producers and consumers.

Beginning with a narrow but instructive framework, internalization occurs whenever property rights for the interdependence are erected and allocated by the appropriate legal system. All other rights and institutions are fixed. Of course, legal inaction regarding the property right arrangements which surround a particular interdependence does not imply that there is no presumptive or putative ownership (Samuels and Schmid, Umbeck). Parties probably have some preconceived notions regarding unvalidated ownership claims. The degree of certainty (or uncertainty) attached to these notions is the foundation of present and future social, political, and legal action by individuals. Externality instigators (i.e., those having the physical power, though possibly not the right, to modify the level of an externality) may clearly create the externality in the absence of prior government action. There is nothing to halt such action by instigators nor to force it.

At a later date a liability or property rule may be newly erected or derived from a more general concept (such as the law of nuisance). Recognition of this possibility may enter the instigator's decision calculus as part of his or her preconceptions regarding the validity of the claim. Uncertainty as to the status of claims will motivate some parties, particularly those with better access to processes of institutional change and/or with positive attitudes regarding the validity or worth of their claims, to pressure the social system for a sanctioned declaration of ownership.

As each externality relationship is born, the process begins: the relationship is tolerated, that is, ownership remains presumptive and unsanctioned, so long as the relation is relatively unimportant. Through an evolutionary process the externality grows or fades in relevance to economic actors as a result of gradual and not so gradual changes in preferences, technology, and relative scarcities. Externalities may diminish into obscurity or evolve into important socioeconomic relationships whereby social declarations of property form the basis of an entire market. Alternatively, trade in the interdependence may not be permitted, and legal constraints can be formed to guarantee the absence of trade (inalienable entitlement). And let us not forget that externalities are resolvable by a large array of social mechanisms other than markets: regulations may be enacted, less formal rules of conduct may be formed, economic incentives may be established, and specific liability rules may be promulgated.

To model these concepts first in the absence of transaction costs, consider an evolving production externality in which the production of good X affects the production of one or more other goods in a manner which is unaccounted for by an existing market. Figure 1 depicts a typical production possibility frontier (PPF) exhibiting social transformation opportunities between good X and the monetary equivalent of all other goods. If the evolving production externality is positive, then the PPF will be shifted outward. If the externality is detrimental, then the shift will be inward. In both cases intercepts A and B will be unaltered. In this way, positive externalities increase the convexity of the PPF , and negative externalities decrease convexity (Baumol).

The impact of the production externality on the potential welfare of consumers is illustrated in utility space. Corresponding to each point on the PPF there is a contract curve depicting Pareto-optimal allocations of all goods between any two consumers. Mapping the contract curve in utility space produces a utility possibility frontier (UPF). The outer limit of all UPF 's (one for each point of the PPF) produces the grand utility possibility frontier ($GUPF$). Through its effect upon the PPF , detrimental production externalities will shift the $GUPF$ inward and beneficial production externalities will shift the $GUPF$ outward. Unlike the PPF , production externalities can alter the endpoints of the $GUPF$. For example, in figure 2 the initiation of a detrimental production externality has shifted the $GUPF$ from CD to $C'D'$. On the other hand, a consumption externality will preserve points C

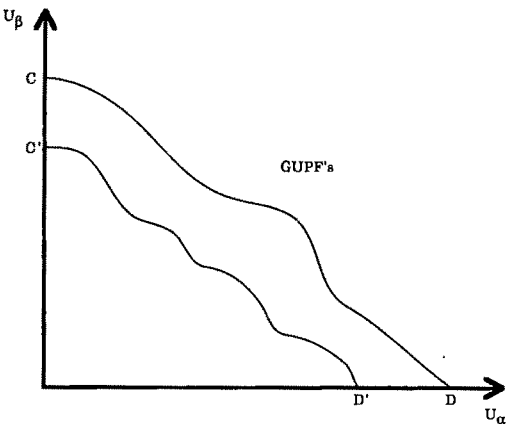


Figure 2. Grand utility possibility frontier shifts for production externalities

and D as endpoints of the $GUPF$. In figure 3, a detrimental consumption externality has produced an inward shift of the $GUPF$. In this latter case, the PPF is unaltered.

The formation of producer \rightarrow consumer externalities may affect the PPF in that consumers may choose to adjust their supplies of production inputs. More important, production at any point along or interior to the PPF will shift the indifference maps of individuals in X - $\$$ space, assuming that the new externality relationship is exogenous to these variables. Contract curves may be affected depending on how indifference maps are shifted, but UPF 's will certainly be shifted. Therefore, the $GUPF$ will be shifted inward or outward in the manner depicted by figure 2.

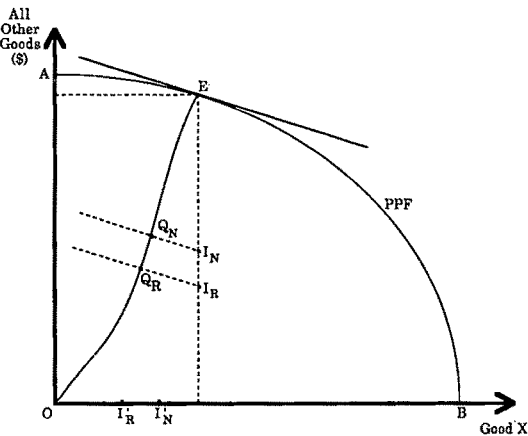


Figure 1. Equilibria for alternative endowments without transaction costs

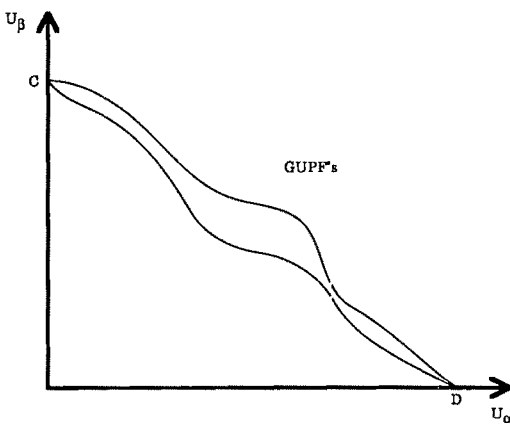


Figure 3. Grand utility possibility frontier shifts for consumption externalities

Suppose that the *PPF* illustrated in figure 1 represents present conditions including the externality. Let us assume that the externality is beneficial and that its instigators (i.e., the producers of X) have the power to produce the externality and that the receptors have no enforceable right to either prevent or force the externality. This initial situation is one of "privilege" favoring instigators and corresponding "no rights" for receptors if legal sanction has not been acquired (Bromley 1989). If legal sanction is obtained, then the situation becomes one of a property "right" for instigators and a corresponding "duty" for receptors (Bromley 1989). Denote this property right as the N (instigator) rule.

The implications of preserving the N rule or adopting its opposite, the R (receptor) rule, has received much attention in the literature.³ This attention has focused almost exclusively on the effect of this choice on the resulting competitive equilibrium and the resulting welfare positions of the economic participants. The following Coasian conclusion is well known and warrants illustration within the above framework: In the absence of private resource constraints, wealth effects, and transaction costs, the Pareto optimality of competitive equilibrium is independent of ownership.

Assume that supply and demand conditions resulting from the initial N rule support a Pareto optimal production equilibrium at point E on the *PPF*. In figure 1 curve OE represents the corresponding contract curve where the economy is assumed to be divisible into two totally homogeneous groups. Group α is composed of the producers of X , all of which have equivalent utility functions. Group β incorporates all others, who are assumed to have equivalent preferences and share equally in their group's net product.

Constructing the Edgeworth box defined by point E , we shall let point O be the origin for the indifference map of group α (the X producers) and point E shall be the origin for group β . After production has been completed, initial endowments will be given by points such as I'_N or I_N . Endowment I'_N will be relevant if X producers have to exchange part of their product to purchase inputs. Endowment I_N will be relevant if X producers own productive factors required by other producers and if the amount of these resources is more than sufficient to acquire the inputs needed for X production.

Within this model, the effect of adopting an R rule to replace the N rule is only to change initial resource endowments. The R rule weakens the welfare position of group α relative to group β regardless of whether the externality interdependence is positive or negative. Thus, the R rule causes new post-production endowments of either I_R or I'_R . Considering the two alternative endowments I_N and I_R , the price vector is given, in both cases, by the tangent to the *PPF* at E . In figure 1 the post-trade distribution of goods is shown for both N and R rules. Q_N and Q_R represent the Pareto optimal allocation of total product for N and R rules, respectively.

Under the above conditions the Pareto optimality of competitive equilibria is independent of ownership. When differential wealth effects are entered into the analysis, the difference in resource endowments will induce income effects which alter the aggregate demands for the two commodities. In this case the production equilibrium for the R rule may be different from point E , the production equilibrium for the N rule. Relative prices will be altered and new welfare positions will be attained. The equilibrium will, nonetheless, be Pareto optimal. Both N and R rules permit trade in the interdependence level, thereby correcting the externality within this highly simplified model.

To summarize, the extant externality represents an irrevocable interdependence among economic agents. If we begin with a situation resembling the N rule, then instituting an R rule constitutes a "reversal" of the active property right. N and R rules, both being acts of internalization through market creation, are beset by the same technical interdependencies among economic actors. Therefore, ownership reversal cannot bring about Pareto improvements in production or utility possibilities. Reversal may easily cause movements along *PPF*'s or *GUPF*'s, but outward shifts in these frontiers will not be possible. The effects of transaction costs represent a major caveat to this rule, a fact to be discussed shortly. For now, we state with confidence the following Coasian postulate.

POSTULATE 1. *Ownership reversal can never improve upon production or utility possibility frontiers in a world without transaction costs.*

Transaction Costs and Social Optimality

It is a trivial extension of presently accepted doctrine to expand the above model into the world

³ At one time N and R rules have been termed L and \bar{L} laws (Mishan), and L' and \bar{L}' are similarly defined "liability rules."

of nonzero transaction costs as follows. *N* and *R* rules differ, in general, with respect to both the total amount of transaction costs that these rules impose and the distribution of these costs. Differences in the distribution of these costs will interact with wealth effects and budget limitations to produce further differences in the resulting competitive equilibria. Distributional issues aside (momentarily), any positive amount of transaction costs will prevent otherwise advantageous transactions from occurring. At this point it becomes appealing to argue that transaction costs prevent competitive equilibria from achieving Pareto optimality. We do not, however, succumb to this temptation.

Transaction costs are real and unavoidable aspects of any economic system. It is not even possible to eliminate transaction costs by prohibiting all trade because such a decree would have to be deliberated and enforced and other institutions would emerge to replace banned markets. The inevitability of transaction costs means that any notion of Pareto optimality is incomplete until transaction costs are incorporated. In what follows we shall set about this task by inquiring how welfare frontiers such as the *PPF* and the *GUPF* are affected by the adoption of either *N* or *R* rules together with their implicit transaction costs.

At the beginning it is crucial to observe a few fundamental points regarding the interdependence which constitutes the general externality to which the *N* and *R* rules are directed. First, the interdependence is produced; it is not an inherent part of the environment but is the result of particular production and/or consumption activities undertaken by man. Second, the production of this interdependence is functionally related to one, several, or many different variables controllable by man. Third, the precise nature of this functional relation may be uncertain; indeed, some of the determining variables may be unknown. Fourth, it is costly to monitor both the interdependence level and its known determinants.

Incorporating Transaction Costs

Let us begin to introduce the effects of transaction costs through example. In the pursuit of killing weeds on my farm, I ameliorate a detrimental externality that I have been perpetrating upon neighboring farmers through the propagation of weed seed. My neighbors derive external benefit from my weed control activities

as I do from similar actions taken by them. The reciprocity of this relationship does not mitigate the externality. I tolerate or take action against my weeds as I see fit, regardless of the actions taken by my neighbors. Each person's neighbors have the duty to receive weed seeds.

Let us contemplate an alternative ownership pattern, an *R* rule, in which my neighbors have an enforceable right to be free of the nuisance induced by my weeds. A pure *R* rule would assign me the duty to manage my farm in such a way as to prohibit the germination of "my" weed seed on any farm but my own. Otherwise, it will be necessary for me to purchase an appropriate amount of weed-seed-propagation rights for my farm—rights which are initially held by the neighbors under the *R* rule.

The terms of the *N* and *R* rules suggested in this example represent technically perfect statements of opposing property rights. On the one hand, I am free to allow my weed seed to germinate on others' land. Under the opposite rule, I must prohibit any weed seed from leaving my property and then germinating. The "perfection" of these rules lies in the fact that the alternative property rights are stated in terms of the actual interdependence, i.e., the amount of weed seed leaving my farm and sprouting on neighbors' farms.

Statement of property rules in any other terms would be technically imperfect. An obvious example would be if an *R* rule prevented me from having weeds on my farm without trading for the rights held by my neighbors. In this case, if I do not purchase rights to have weeds on my farm, then I must eliminate all weeds. No other method of controlling the spread and germination of weed seed is admitted. Clearly, the propagation of weed seed is related to the amount of weeds, but the relationship is not complete. For example, if I trim my weeds frequently, the propagation of seed will be minimal. I could chemically retard the formation of seed without killing the weeds. Or, at the extreme, my weeds may be of a hybrid, infertile variety. A property rule which is stated in terms of a surrogate measure instead of the actual interdependence will not recognize these and other possibilities. This situation applies to all externalities but to differing degrees.

Establishing a workable *R* rule favoring one's neighbors in the weed seed scenario requires that an index/measure be formulated. At least four alternatives exist for an *R* rule favoring one's neighbors. We could state that each landowner (a) may not allow his/her weed seed to ger-

minate on others' land, (b) may not allow his/her weed seed to move to others' land, (c) may not possess reproducing weeds, or (d) may not grow any weeds. Trade can occur following the adoption of any of these *R* rules, but only option (a) embraces the actual interdependence. Options (b)–(d) employ surrogates for the interdependence and therefore embody misconceptions regarding the interdependence. This is a problem of transaction costs because surrogates are otherwise unnecessary.

The essential point is that technically imperfect property rules do not recognize, by definition, certain externality control processes. If the property rule ignores a control mechanism which would have been employed under costless information, then the rule results in a *PPF* and/or a *GUPF* which is interior to the idealistic frontiers which ignored transaction costs. In a world devoid of transaction costs, imperfect formulations of property rules will be unimportant to efficiency. Without transaction costs, information shall be perfectly and universally available to all. Actual interdependence levels will be perfectly and costlessly monitorable, and precise relationships between externality generation and externality determinants will be known. In the absence of such idealistic conditions we have:

POSTULATE 2. If a particular property rule is technically imperfect, then the PPF and GUPF will be shifted inwards even though transaction costs are otherwise zero.

According to postulate 2, the effects of an imperfect property rule for externality "correction" cannot be fully overridden, and this restrains production and/or utility possibilities. While market activity with an imperfect rule presumably will bring about some mutually beneficial transactions, some beneficial transactions will be avoided and some bad transactions will occur because of incorrect judgments regarding the nature of interdependencies.

On the other hand, use of surrogate measures in property rules is a justifiable response to transaction costs. A technically perfect property rule can be informationally demanding in practice and thereby incur high transaction costs. Of the four identified options for defining a weed seed *R* rule, only option (a) is technically perfect. It is also economically onerous, for it is difficult to imagine the costliness of a monitoring network capable of tracking the origin, transport, and growth of weed seeds. Clearly, an economically rational response to transaction

costs may be to select interdependence surrogates for property rule formulation. Why else would we buy and sell oranges by the pound when we truly desire juice and pulp of good quality (Barzel)? Why do western water rights specify acre-feet diverted rather than consumed? How closely does tons per acre per year really represent complex soil erosion interdependencies? Surrogate measures of externality interdependencies are important, and postulate 2 is therefore relevant.

The arguments presented thus far have not addressed the nature of the respective welfare frontiers for *N* and *R* rules that utilize equivalent indices. While transaction costs can exhibit some similarities for opposing *N* and *R* rules employing the same index, the effect is likely to be unique. The following postulate offers one reason.

POSTULATE 3. Transaction costs increase with the distance between initial endowments and final (post-trade) allocations.

The most important implication of this postulate is that transaction costs convey inertia to initially erected property rights. Because market participants are reluctant to incur transaction costs, the restraint to trade caused by the existence of transaction costs implies that post-trade allocations of rights will bear some resemblance to initial endowments.⁴ The extent of the resemblance depends on both the magnitude and distribution of transaction costs: greater transaction costs giving rise to a greater resemblance between pre-trade and post-trade allocations. The distribution of these costs is also pertinent. For example, to the degree that transaction costs are borne by individuals (or entities) who are not directly involved in the interdependence (e.g., a government authority), the inertia of initial endowments will be lessened.

Once a particular property rule and the mechanism for enforcing this rule are adopted, the amount and distribution of transaction costs are more clearly assigned (though still unknown). That is to say, the burden of information generation is more apparent. As a consequence of adopting a specific rule, traditional production and utility possibility frontiers must be shifted inward to reflect the amount of resources which are consumed for the production of information

⁴Randall (1983) has observed this consequence of transaction costs and has also stated that the income effects of alternate property rules cause similar differences in post-trade allocations.

and lost because of imperfect decisions made under imperfect information. That is, the traditional concepts of *PPF* and *GUPF* do not incorporate the very real costs imposed by transaction costs. Once property institutions are known, adjustments to welfare frontiers within our theoretical constructs can and should be performed to account for these costs.

POSTULATE 4. *Once transaction costs are admitted, different property rules give rise to different welfare frontiers.*

In figure 4 the *PPF* relating to a hypothetical *R* rule is drawn as the locus *PPF_R* which lies strictly interior to the original, idealistic frontier. For generality *PPF_R* lies strictly interior to the old frontier to indicate that information generation will occur even in the absence of transactions. Transaction costs incurred by the public sector would also be incorporated in determining *PPF_R*. Property right inertia due to the inherent transaction costs of the *R* rule implies a proportionately greater inward shift along the lower end of the *PPF* (assuming that the *R* rule was imposed for a detrimental externality).

Similarly, *PPF_N* represents production possibilities for the *N* rule. Presumably, inertia under the *N* rule will favor greater production possibilities for *X* in comparison to the *R* rule. Under the *R* rule, larger amounts of bargaining and monitoring activities must be pursued in order to have a larger production of *X*, that is, a greater number of rights must be exchanged.

Each of these production frontiers relates to a specific *GUPF* which also incorporates trans-

action costs. It is not, however, true that intersecting production possibility frontiers produce intersecting grand utility possibility frontiers. It is therefore feasible for either *GUPF_R* or *GUPF_N* to be everywhere superior.

Regardless of the relative positions of the two pairs of welfare frontiers, we might impose social indifference curves of the form

$$W = W(U_{\alpha}, U_{\beta})$$

on the alternative *GUPF*'s, as in figure 5, in order to determine whether the *N* or *R* rule is socially preferred. Equivalently, as in figure 6, welfare indifference curves of the Bergson variety,

$$B = B(\$, X),$$

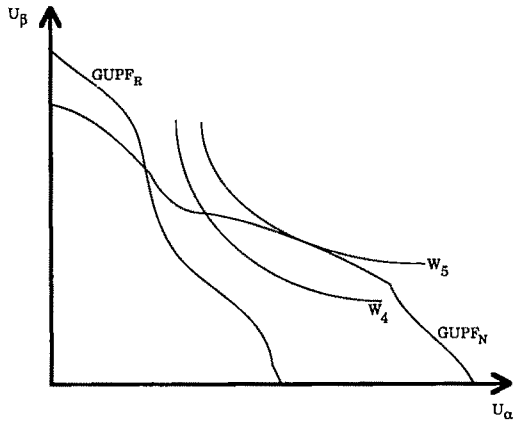


Figure 5. Social optimization in utility space

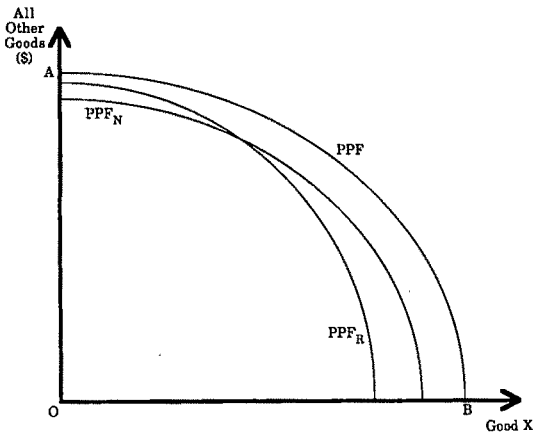


Figure 4. Production possibilities for alternative endowments with transaction costs

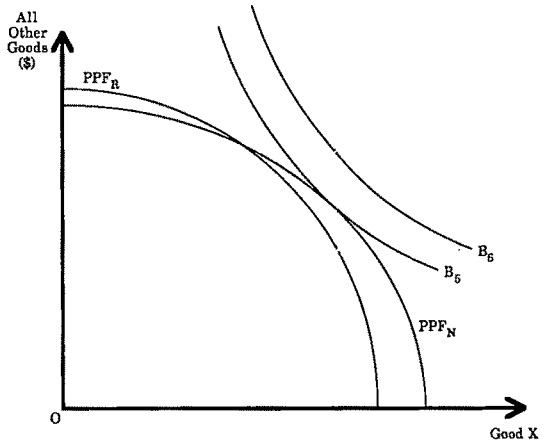


Figure 6. Social optimization in goods space

could be drawn together with the two alternative *PPF*'s to illustrate the socially optimal property rule. Either approach is flawed.

Choice between Opposing Property Rules

While the imposition of social indifference curves is a nice theoretical device for illustrating the "optimal" choice of property rule with transaction costs, it is neither realistic nor helpful. Beyond the obvious acknowledgement that we do not possess a social welfare function, these devices treat every point on the welfare frontiers as being attainable even though each frontier corresponds to a particular distribution of the property rule in question.

The welfare frontiers in effect for the two distinct property rules can become determinant only by incorporating the burden of information costs. This, in turn, requires distribution of the property right under examination. However, this same fixity tends to define the resulting competitive equilibrium by declaring an initial endowment. Each such endowment will presumably support a unique competitive equilibrium which is represented by a single point on the appropriate welfare frontier.⁵ Because true welfare frontiers must incorporate transaction costs and because the inclusion of transaction costs requires property right specification which then fixes the competitive equilibrium, the concepts of Pareto optimality and competitive equilibrium appear to come perilously close to being definitionally identical. Randall (1983) has employed this closeness to argue that "externality" and "Pareto relevance" are vacuous terms. This argument is feasible only when nonmarket institutions are assumed inferior and no Pareto-inefficient definitions of property rights are envisioned.

Focusing now on the correct interpretation of the choice between polar *N* and *R* rules, it is clear that the two alternatives present distinct welfare frontiers. Furthermore, only one point

on each of these frontiers, the competitive equilibrium, is relevant to the decision. In figures 7 and 8 the diagrammatic illustration of a newly emerged production externality is continued.⁶ Regardless of whether the new interdependence is positive or negative, the *N* rule favors the producers of the externality, in this case the producers of *X*. As depicted in figure 7, the *N* rule results in an initial endowment given by I_N , an output vector given by F , and an output distribution of Q_N . The determination of these three points is simultaneous; even I_N is dependent on F since I_N represents post-production claims on net output.⁷ Because property right inertia brought on by transaction costs will favor greater *X* production under an *N* rule than an *R* rule, point F lies to the left of point G . Post-production endowments and output distribution are similarly depicted for the *R* rule.

Contract curves OF and OG in figure 7 define utility possibility frontiers in utility space where α and β are the groups of completely homoge-

⁶Production-consumption and pure consumption externalities can also be illustrated in this framework. Where production activities have externality impacts on utility functions, indifference mappings (and contract curves and utility possibility frontiers) will depend on equilibrium production. Where production opportunities are affected by consumption activities, production possibility frontiers will be jointly interrelated to actual consumption. Nonmarket interrelationships between consumers will likewise shift indifference maps and utility possibility frontiers. In all of these situations *N* and *R* rules will necessarily induce disparate utility possibility frontiers. Moreover, since wealth effects may generally occur and transaction costs may vary, different production possibility frontiers may result from *N* and *R* rules.

⁷As demonstrated with figure 1, the placement of I_N is rather arbitrary. Its exact location is not important for the purposes of the present discussion.

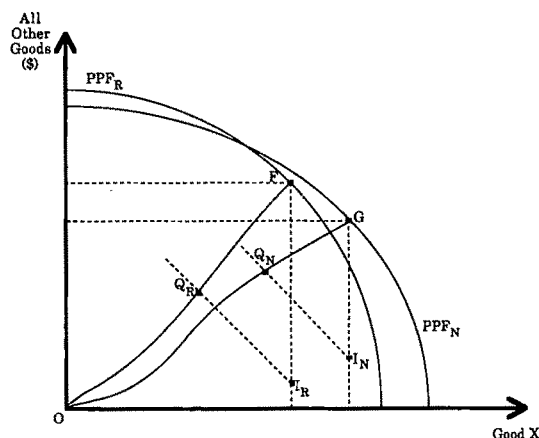


Figure 7. Equilibria for alternative endowments with transaction costs

⁵Lump-sum transfers of other resources could be mistakenly viewed as making viable equilibria out of other points along $GUPF_N$ or $GUPF_R$. This is an oversimplification that underestimates the difficulty with which redistributive policy is performed. It is one thing to design an externality-addressing property rule of the sort under discussion here and quite another to pair such a rule with an unspecified redistributive policy. Such a policy would be accompanied by its own unique burden of transaction costs. In the face of these costs neither $GUPF_N$ nor $GUPF_R$ is relevant. Another frontier recognizing the particulars of the chosen transfer policy is needed. Each property rule or each pairing of a property rule and a lump-sum transfer policy will be associated with a unique *UPF*. Subsequent action by individuals will lead to a point on this frontier.

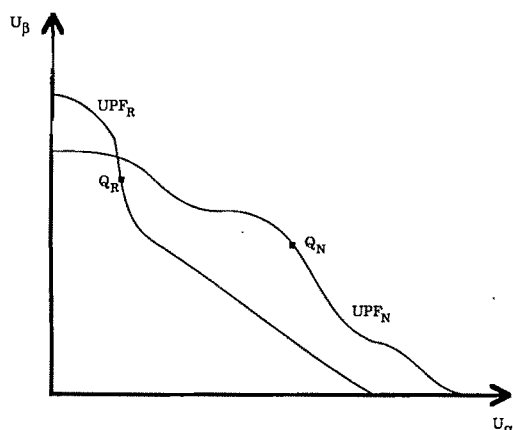


Figure 8. Equilibria with transaction costs in utility space

nous X -producers and all others, respectively.⁸ These welfare frontiers are illustrated in figure 8 as are the equilibrium utility levels resulting from human action within the whole system of preferences, technology, and institutions. Points Q_R and Q_N completely summarize the welfare implications of property right choice. There are no conditions to prevent one of these equilibria from lying within the frontier defined by the other property right (as depicted in fig. 8). Nor is this relevant. The choice between property rules is a choice between the equilibrium outcomes rather than between the frontiers on which the outcomes lie. Each welfare frontier is not completely feasible. Economic activity cannot lead to an arbitrary point of the frontier.

POSTULATE 5. *The incorporation of transaction costs required for realistic welfare frontiers has the side effect of rendering irrelevant all but one point on the welfare frontier. The choice between opposing N and R rules is a choice between two and only two economic states.*

Therefore, the choice between N and R rules is reduced to a choice between Q_N and Q_R . The welfare frontiers are inconsequential, and the only choice involves whether group α or β is to be favored. Moreover, the choice itself is likely dominated by pre-existing social institutions constituting the fabric of mutual coercion. In specific instances these institutions may rigidly dictate which property rule (and which group)

is to be ascendant. Of course, to the extent that flexibility in property right choice remains, we can expect each group to exploit its noneconomic opportunities for property right capture.⁹

Because of these facts, it is rarely possible for economists to interpret property rule reversals as economically inefficient. Regardless of the relative positions of welfare frontiers, the alternative economic outcomes will be Pareto non-comparable as in figure 8. Only in rare cases where Q_N is preferred to Q_R (or vice versa) by both α and β can a clear position be taken.

Institutional Choice

By utilizing two different applications of the same institution, the preceding discussion has demonstrated the dependence of economic efficiency on institutional choice. Were it not for transaction costs, postulate 1 would apply, and Pareto optimality would be divorced from the selection of institution. It should be noted that the Pareto optimality of competitive equilibria has not been investigated here (as in the Coase theorem) but has been taken for granted. Previous sections have considered the ramifications of incorporating transaction costs into the determination of Pareto optimality. By making transaction costs endogenous, the choice between N and R rules becomes a choice between two different economic states, both being Pareto optimal. Production and utility possibility frontiers are irrelevant because they are infeasible. Each of these property rules is capable of supporting a single Pareto optimum.

It is now appropriate to enrich this analysis with a consideration of internalizing institutions other than N and R property rules. Many such institutions are important, but an exhaustive listing of these would be tedious. Even so, considerable insight can be achieved by inspecting some of the more noteworthy alternatives. First among these is the multitude of property rules lying along the spectrum whose endpoints are defined by N and R rules.

Other Property Rules

Between the two extremes represented by N and R rules, there is a large, indeed infinite, number

⁸The presumed uniqueness of the competitive equilibrium implies that only one point on the production possibility frontier (and, therefore, only one contract curve) is relevant. Thus, a single utility possibility frontier, not the grand utility possibility frontier, is the appropriate concept relating to a specific property rule.

⁹As Bromley has observed, it is the incidence of gains rather than the opportunity to obtain net social gains that motivates institutional change (1989, p. 240).

of alternative property rule specifications. These choices differ with respect to the endowed level of interdependence that instigators can/must generate and, simultaneously, the endowed level of interdependence that receptors must tolerate/enjoy. Expanding upon Bromley (1989), it could be said that we are now exploring alternate linear combinations of rights and duties. For example, property rules could vest farmers with (a) the right to grow 100 weeds per acre and the duty not to exceed this amount, (b) the right to generate two tons of soil loss per acre per year and the duty not to lose more soil than this, (c) the duty to produce 200 pounds of oranges each year and the right to output in excess of this amount, etc. Transfer of these endowments is allowed and is the basis of market activity. In the case of the typical marketed commodity, manufacturers are empowered to produce and withhold the commodity. Potential consumers must tolerate this authority unless they purchase a portion of these rights.

In the case of a positive externality a property rule can be adopted which forces each instigator to produce y units of the interdependence. At the same time such a rule will specify the initial allocation of these y units (assuming rivalness) among the interdependence benefactors. The quantity y can be any positive number, thereby illustrating the infinite possibilities for alternative property rules. Clearly, if y is identically zero, then we have an N rule. As y gets large, the property right trends towards becoming an R rule. This indicates an area where the earlier discussion was not definitive. In the case of a beneficial interdependence the magnitude of y needed to erect an R rule is ambiguous.¹⁰

Consideration of property rights which are intermediate to N and R rules or which employ alternative surrogate indices for the true interdependence simply generalizes earlier discussion. Some of the more important conclusions bear restatement under these more general conditions.

- Each alternative property rule will exact its own unique magnitude and distribution of transaction costs. In each case ensuing transactions will be vectored by the current information base regarding known and economically measurable determinants of the externality interdependence.
- Property rule inertia induced by transaction

costs will favor the initial distribution. In this way different property rules imply different distributions of attainable economic welfare.

- Each distinct property rule will be associated with a single attainable point in utility space (postulate 4). In general, many of these points will be Pareto noncomparable.

Using the same techniques employed to obtain figures 7 and 8, the attainable levels of welfare associated with alternative property rules can be shown in utility space. The outer limit of these points depicts the choices inherent to the choice among alternative property rules (fig. 9). This locus (possibly discontinuous) might be termed the Property Rule UPF . When property rules represent the only institutional variant for addressing an externality relationship, this frontier fully depicts the scope and implications of institutional choice. Every point on this frontier is attainable in the sense that actual transaction costs are endogenous.

The welfare frontier depicted in figure 9 must not be confused with the traditional concepts of UPF 's or $GUPF$'s. UPF 's are utility space mappings of contract curves, and a $GUPF$ is the outer limit of an infinite number of UPF 's (one for each point on the PPF). The property rule UPF in figure 9 is composed of single points from a particular class of UPF 's incorporating transaction costs. The Property Rule UPF is not an outer limit of these UPF 's, but it is the outer limit of competitive equilibria lying on the UPF 's. The property rule for a single externality interdependence is variant along the welfare frontier depicted in figure 9; all other institutions are fixed. Social constraints which might narrow the feasible set of property rules for this

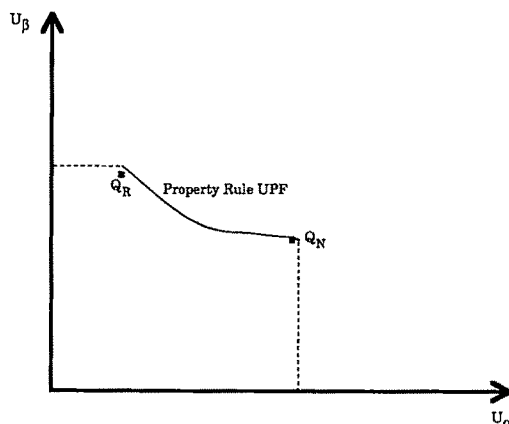


Figure 9. Property rule utility possibility frontier

¹⁰Similarly, in the case of a detrimental externality y would denote the amount of the interdependence that instigators are allowed to produce. Under these conditions $y = 0$ implies an R rule, and the size of y necessary to achieve a pure N rule is indeterminate.

interdependence have been disregarded in the construction of this frontier.

Liability Rules

The welfare frontier originating from the preceding discussion lays the basis for similar analysis involving broader dimensions of institutional choice. Certainly, the analysis should be extended to include liability rules. It has been demonstrated that liability rules differ dynamically from property rules (Buchanan and Faith, Frech). The static nature of the economic tools being employed here prevents immediate attention to these dynamic elements. There appears to be little opportunity to usefully depict the effects of today's institutions on tomorrow's welfare frontiers except by adding a time dimension to the welfare frontiers, and this extension is not pursued here.

Liability rules also differ from property rules in terms of transaction costs and the degree of afforded protection. Liability rules offer less protection to endowed individuals than do property rules. Therefore, the liability rule possesses less market value. Transaction costs are distributed quite differently when comparing liability rules to property rules (Bromley 1978), and, since the *ex post* determination of value in the case of liability rules requires an expensive judicial system, a large portion of transaction costs are borne by third parties. Like property rules, alternative surrogates are available and liability rules are continuously assignable, and, therefore, a Liability Rule *UPF* similar to that in figure 9 is presumably applicable.

Government Externality Policy

The final major category of externality-addressing institutions to be separately considered here pertains to various forms of government policy. Types of government policy traditionally discussed by economists include quantity-guided controls such as prohibition, standards, and regulation in general, as well as price-guided controls such as taxes, charges, subsidies, cost-sharing, and various other economic incentives. Perhaps the only characteristic serving to unify this class of institutions is that they are all modes of nonmarket activity. However, economic incentives are usually characterized as decentralized while regulation is quite centralized (Schultze), and a great deal of literature is devoted to contrasting these two subgroups.

Like the variety of property institutions which address externality interrelationships, the purpose of these alternative forms of nonmarket activity is to assist in establishing social and economic structure. The use of surrogate measures is also relevant. Market and nonmarket institutions are quite similar in that neither avoids transaction costs and in that reversal is a relevant approach for contrasting polar images of the same institution. For example, in the case of beneficial externality, a regulation can conceivably state that a producer of the externality need not produce the interdependence at any level or, alternatively, that a producer must produce y units of the interdependence. Because these are regulations, trade in these units between instigators and receptors is not possible in either case. If the externality takes on aspects of a rival good in that its use is consumptive, then the regulation might also state a distribution of these y units among recipients.

In this way, transaction costs pertaining to the direct relationships between instigators and receptors are greatly lessened. Instead, transaction costs are incurred via the relationships between instigators and government, on the one hand, and between government and receptors on the other. As with market exchange processes, transaction costs can be quite substantial, and polar regulations or polar incentives will differ greatly in the magnitude and distribution of transaction costs. Because of the existence of transaction costs, perfect enforcement of regulations is not attractive, and, depending on the magnitude of administrative transaction costs, resulting economic activity will stray from promulgated regulations. Still, as is the case with property rules, resulting economic activity will bear some resemblance to established regulations and will be sensitive to regulatory choice.

It must therefore be clear that alternative forms of government policy or alternative specifications of the same policy will relate to different levels of attainable welfare. Technology is unaffected by available policies in this static setting; but, once a policy is chosen, transaction costs influence the outcomes which result in attained welfare. Again, if the concept of economic efficiency truly relates to attainable levels of welfare, then the concept cannot be independent of policy alternatives. Different policy actions induce different levels and distributions of transaction costs, and it is not accurate to say that one particular policy is economically efficient without accounting for these differences.

It comes as no surprise at this point that each nonmarket (policy) alternative is to be associ-

ated with a single economic outcome in terms of production, consumption, and utility levels. This outcome is selected by the whole of individual and collective action from transaction-cost-endogeneous welfare frontiers. Moreover, the results of many of these policies will not be Pareto comparable, and the locus of outermost production or utility levels can be used as an efficiency frontier for nonmarket action.

Institutional Welfare Frontiers

Because Pareto optimality cannot be divorced from institutional alternatives for a given externality, any welfare frontier depicting the range of efficient economic states must consider a variety of institutions. Institutional alternatives are endogenous to the construction of true welfare frontiers. It is evident at this point that the revised construction of welfare frontiers, elucidated earlier in the context of variable property rules, can be readily expanded to incorporate a much broader range of institutions.

Implementation of different property rules, liability rules, regulation, incentives, customs and behavioral standards, and other nonmarket devices are construed as separate institutions with distinct economic consequences. The inherent transaction costs of each specification of each institution will produce an institutionally specific production possibility frontier and utility possibility frontier, but individual action within the system of constraints and incentives will succeed in isolating a single point on each of these frontiers. Other points on these frontiers are not feasible (in the sense that they will not be chosen) and, therefore, are not entertainable as efficient. Various economic states then correspond to various institutions. Where these states are Pareto comparable, specific choices can be identified as inefficient. Minimally, there are property rule, liability rule, and nonmarket UPFs applicable for each externality. Superimposing these welfare frontiers and forming the outer limit to eliminate all inefficient alternatives will reduce the institutional choice set to a Pareto non-comparable group. This group is representable by a curve or surface in utility space which shall be called the Institutional UPF (*IUPF*) (not illustrated). The modifier "grand" is unnecessary here, but the frontier is, indeed, grand in the sense that it embodies the full spectrum of internalizing actions available for a given interdependence. All forms of ownership, protection, nonmarket policy, and, in general, all aspects of mutual coercion are endogenous to

the *IUPF*. Grand utility possibility frontiers of earlier literature only incorporated resource and technological constraints. The *IUPF* also accounts for the range of humanly imposed constraints available to order human interaction.

An institutional *PPF* can also be illustrated in output space; each point on this efficiency locus corresponds directly with a point on the *IUPF* and a particular institutional specification. However, the *IUPF* is generally a more fundamental and interesting concept because it highlights resultant utility levels. While variable institution welfare frontiers definitely connote abstract concepts, they are no more abstract than traditional welfare diagrammatics and are far more useful as a paradigm for institutional choice.

Implications for Portraying Institutional Choice

Within the framework provided by variable institution welfare frontiers, shifts in property assignments and nonmarket policies are best depicted as movements within the region bounded by the *IUPF*. This pertains to the choice of new or different institutions as well as incremental changes in a certain type of institution. Pareto-relevant externalities are properly characterized as interdependencies currently being addressed by market/nonmarket structures associated with economic states that are strictly interior to the *IUPF*. Pareto-irrelevant externalities are those relationships being accommodated by institutions which place the economy on the *IUPF*.

One approach to "remedying" an externality is judged to be better than another on the basis of the Pareto criterion, and, in this way, the *IUPF* leads to an institutionally unbiased comparison of alternative choices. The potential usefulness of variable institution welfare frontiers as a theoretical concept stems from this unbiasedness. In practice, it is unacceptable to define externality as an "interdependence without a market." An externality is merely an interdependence. If an externality is Pareto-relevant (we are not on the *IUPF*), then some sort of policy is needed, but that policy is not necessarily a market.

Most important, these frontiers are not idealistic in that an unachievable norm for efficiency is maintained. Variable transaction costs and the effects of these transaction costs are endogenous. Social action for externality amelioration is therefore justified if current institutions do not place the economy on the *IUPF* or if current institutions do place the economy on this fron-

tier but with undesirable distributional consequences.

One purpose of this theory is to give explicit recognition to the fact that externalities do not emerge into institutional voids but are always being addressed by some, perhaps very subtle, coercive structure. In this sense externality correction, resolution, and internalization are all vague terms unless they refer to institutions which, if adopted, would move society to the *IUPF*. Because a proper analysis incorporating transaction costs has never been performed to investigate the global efficiency of a prospective institution, the applicability of such terms is highly questionable in all but conceptual work. Moreover, the empirical difficulties to be encountered in such a rich analysis imply that the chances of ever satisfying this requirement are quite remote.

Finally, it is impossible to appeal to economic efficiency (in the Paretian sense) as a norm to guide institutional choice. Taking the *IUPF* and dimensioning it into hyperspace to accommodate a utility axis for every individual has the effect of grossly multiplying the number of Pareto noncomparable (and therefore efficient) institutions. The economic analysis of social options intended to accommodate a given externality interdependence will rarely be able to consider a wide range of alternatives. Agenda formation within decision-making and research processes almost certainly excludes many institutional alternatives because of the limited scope of these processes and our limited abilities to understand and characterize many alternatives. Institutional bias may always be present because selection is from a subset of the actually available institutions.

[Received February 1990; final revision received September 1990.]

References

- Barzel, Yoram. "Measurement Costs and the Organization of Markets." *J. Law and Econ.* 25(1982):27-48.
- Bator, Francis M. "The Simple Analytics of Welfare Maximization." *Amer. Econ. Rev.* 47(1957):22-59.
- Baumol, William J. "External Economics and Second-Order Optimality Conditions." *Amer. Econ. Rev.* 54(1964):358-72.
- Baumol, William J., and Wallace E. Oates. *The Theory of Environmental Policy*. Englewood Cliffs NJ: Prentice-Hall, 1975.
- Bromley, Daniel W. *Economic Interests and Institutions*. New York: Basil Blackwell, 1989.
- . "Markets and Externalities." *Natural Resource Economics*, ed. D. W. Bromley, pp. 37-68. Boston: Kluwer Academic Publishers, 1986.
- . "Property Rules, Liability Rules, and Environmental Economics." *J. Econ. Issues* 12(1978):43-60.
- Buchanan, James M., and Roger L. Faith. "Entrepreneurship and the Internalization of Externalities." *J. Law and Econ.* 24(1981):95-111.
- Calabresi, Guido, and A. Douglas Melamed. "Property Rules, Liability Rules, and Inalienability: One View of the Cathedral." *Harvard Law Rev.* 85(1972):1089-1128.
- Dahlman, Carl J. "The Problem of Externality." *J. Law and Econ.* 22(1979):141-62.
- Demsetz, Harold. "The Exchange and Enforcement of Property Rights." *J. Law and Econ.* 7(1964):11-26.
- . "Toward a Theory of Property Rights." *Amer. Econ. Rev.* 57(May 1967):347-59.
- Frech, H. E., III. "The Extended Coase Theorem and Long-Run Equilibrium: The Nonequivalence of Liability Rules and Property Rights." *Econ. Inquiry* 17(1979):254-68.
- Heller, Walter I., and David A. Starrett. "On the Nature of Externalities." *Theory and Measurement of Economic Externalities*, ed. Steven A. Y. Lin, pp. 9-22. New York: Academic Press, 1976.
- Macpherson, C. B. "The Meaning of Property." *Property: Mainstream and Critical Positions*, ed. C. B. Macpherson, pp. 1-13. Toronto: University of Toronto Press, 1978.
- Mishan, E. J. "The Postwar Literature on Externalities: An Interpretative Essay." *J. Econ. Lit.* 9(1971):1-28.
- Randall, Alan. "Market Solutions to Externality Problems: Theory and Practice." *Amer. J. Agr. Econ.* 54(1972):175-83.
- . "The Problem of Market Failure." *Nat. Resour. J.* 23(1983):131-48.
- Samuels, Warren J., and A. Allan Schmid. "Polluters' Profits and Political Response: The Dynamics of Rights Creation." *Pub. Choice* 28(1978):99-106.
- Schultze, Charles L. *The Public Use of the Private Interest*. Washington DC: The Brookings Institution, 1977.
- Umbeck, John. "Might Makes Right: A Theory of the Formation and Initial Distribution of Property Rights." *Econ. Inquiry* 19(1981):38-58.

Factor Demands in the U.S. Food-Manufacturing Industry

Kuo S. Huang

This paper analyzes the demand for labor, capital, and energy in the U.S. food-manufacturing industry using Allen and Morishima elasticities of substitution. The demand for capital is more elastic than for labor and energy, and these production factors are substitutable, especially between capital and labor.

Key words: conditional factor demands, elasticity of substitution.

Some U.S. food-processing technologies require heavy use of energy and capital equipment, while labor costs relative to value of shipments are smaller than for other types of manufacturing. In past years, the food manufacturers faced a steady increase in the price ratio between labor and capital and a sharp increase in the price of fuel and energy during the 1970s but a continued decrease since 1981. To make decisions in resource allocation, food manufacturers need information on the nature of industrial demand for factor inputs.

Numerous studies have analyzed demand for production factors in various sectors of an economy. Examples include Berndt and Christensen, Berndt and Wood, and Fuss in the general manufacturing sector, and Binswanger, Ray, Lopez, and Shoemaker in the agricultural sector. However, little attention has been given to factor demands in the U.S. food-manufacturing industry. The objective of this study is to analyze the demand for nonfood inputs such as labor, capital, and energy in U.S. food manufacturing. Agricultural or marine raw products and packaging and containers are not considered, partly because of little perceived interdependent relationships between materials and nonfood inputs and partly because of difficulty in measuring quantity and price.

In addition to the Allen partial elasticity of substitution (*AES*), a major focus is on measuring the Morishima elasticity of substitution (*MES*) to explain factor demands and their interdependent relationships. Because *MES* is not frequently used by applied economists, a brief ex-

planation of *MES* and its linkage to cost function is given followed by an application to the U.S. food-manufacturing industry.

Conceptual Framework

Consider a production technology $F(X)$ for a vector of n -factors X . The corresponding cost function is defined as the minimum cost of attaining product Q at a vector of factor prices W as

$$(1) \quad C(W, Q) = \min_x [W'X; F(X) = Q].$$

This cost function is concave and linear homogenous in W .

Denoting $C_i(W, Q)$ and $C_{ij}(W, Q)$, respectively, as the first- and second-order partial derivatives of the cost function with respect to factor prices, one can apply Shephard's lemma and derive a conditional factor demand for the i th factor X_i^* as a function of W and Q :

$$(2) \quad X_i^*(W, Q) = C_i(W, Q).$$

This conditional factor demand function is homogenous of degree zero in factor prices. Furthermore, the (constant-output) cross-price elasticity, E_{ij} , for i th factor with respect to j th factor price is then obtained as

$$(3) \quad E_{ij} = W_j C_{ij}(W, Q) / C_i(W, Q).$$

Hicks defined the elasticity of substitution for a two-factor production function as the ratio of factors in response to a change in their relative prices. Later Allen (p. 504) extended the definition to account for the adjustments for more than two factors and defined "partial elasticity

Kuo S. Huang is an agricultural economist with the Economic Research Service, U.S. Department of Agriculture.

The author wishes to thank Lester Myers and *Journal* reviewers for helpful comments.

of substitution" between i th and j th factors as below:

$$(4) \quad AES_{ij} = \left\{ \left[\sum_{i=1}^n X_i^*(W, Q) F_i \right] / [X_i^*(W, Q) X_j^*(W, Q)] \right\} (F_{ij}^*/F),$$

where F_i is the marginal product of i th factor, F is a determinant of the Hessian matrix (its element denoted as F_{ij}) bordered by marginal products, and F_{ij}^* is a determinant of the cofactor of F_{ij} in the matrix of F .

Although AES has been widely used, its applicability to the demand for production factors is rather limited. The definition of AES deviates from Hicks' definition for two-factor production and does not explain factor substitution explicitly. Besides, AES is not a measure of the curvature of the isoquant, it provides no information about relative factor shares, and it cannot be interpreted as the marginal rate of substitution (Blackorby and Russell).

An alternative measure of factor substitution known as the Morishima elasticity of substitution is defined as a logarithmic derivative of a quantity ratio in factors with respect to a ratio of its factor prices:

$$(5) \quad MES_{ij} = -\partial \ln[X_i^*(W, Q)/X_j^*(W, Q)] / \partial \ln[W_i/W_j].$$

MES measures the percentage change in the ratio of a pair of factors in response to a change in their relative prices. It is a natural generalization of the Hicksian two-variable elasticity.

Because the conditional factor demand function (2) is homogenous of degree zero in factor prices, the demand function is invariant by dividing through the prices with a particular price, say W_j ; that is,

$$(6) \quad X_i^*(W, Q) = X_i(W_1/W_j, \dots, W_{j-1}/W_j, W_{j+1}/W_j, \dots, W_n/W_j, Q).$$

The differentiation can be carried through the chain rule by differentiating the $X_i(\cdot)$ directly with respect to the variable expressed in terms of W_i/W_j . The result is a workable form of MES expressed as

$$(7) \quad MES_{ij} = E_{ji} - E_{ii}.$$

This derivation is more convenient and straightforward than the expression derived in Blackorby and Russell.

According to equation (7), the effect of a variation in the factor price ratio W_i/W_j can be divided into two components: (a) the effect on $X_j^*(W, Q)$ given by the cross-price elasticity E_{ji} , and (b) the effect on $X_i^*(W, Q)$ given by E_{ii} . One property of MES is asymmetry in that the effects of change in W_i/W_j and W_j/W_i upon their corresponding adjustments to the ratio of factor demands need not be the same. MES can also provide complete comparative static information about relative factor cost shares in response to a change in factor prices expressed as

$$(8) \quad \frac{\partial \ln[W_i X_i^*(W, Q)/W_j X_j^*(W, Q)]}{\partial \ln[W_i/W_j]} = 1 - MES_{ij}.$$

The relative cost share is decreasing (increasing) if the MES is greater (less) than one.

Thus far, factor demand relationships are expressed in terms of an unknown cost function. Some empirical applications use a translog cost function expressed as

$$(9) \quad \ln C = \sum_{i=1}^n \alpha_i \ln W_i + 1/2 \sum_{i=1}^n \sum_{j=1}^n \beta_{ij} \ln W_i \ln W_j + \sum_{i=1}^n \beta_{iq} \ln Q \ln W_i + \alpha_q \ln Q + 1/2 \beta_{qq} (\ln Q)^2 + \sum_{i=1}^n \delta_i \ln W_i T + \delta_q \ln QT + \mu_1 T + \mu_2 T^2,$$

where variables C is cost, Q is production, W_i is the i th factor price, and T is a time trend to represent the level of technical progress.

To ease potential estimation problems such as lack of degrees of freedom, the cost function is not estimated directly; rather, a set of the factor cost share equations is estimated. A typical cost share equation for i th factor, say S_i , derived by applying Shephard's lemma to the cost function is expressed as

$$(10) \quad S_i = \alpha_i + \sum_{j=1}^n \beta_{ij} \ln W_j + \beta_{iq} \ln Q + \delta_i T \quad i = 1, 2, \dots, n.$$

Moreover, in order to satisfy the properties of neoclassical production theory and the adding-up property of cost shares, the following parameter restrictions are required:

$$\sum_{i=1}^n \alpha_i = 1, \beta_{ij} = \beta_{ji} (i, j = 1, 2, \dots, n), \text{ and} \\ \sum_{j=1}^n \beta_{ij} = \sum_{i=1}^n \beta_{ij} = \sum_{i=1}^n \beta_{iq} = \sum_{i=1}^n \delta_i = 0.$$

Given the estimated cost parameters, one can derive *MES*, *AES*, and the price elasticities of conditional factor demands as follows:

$$(11) \quad MES_{ij} = (\beta_{ij} + S_i S_j) / S_j \\ - (\beta_{ii} + S_i^2 - S_i) / S_i \quad (i \neq j),$$

$$(12) \quad AES_{ii} = (\beta_{ii} + S_i^2 - S_i) / S_i^2,$$

$$(13) \quad AES_{ij} = (\beta_{ij} + S_i S_j) / S_i S_j \quad (i \neq j),$$

$$(14) \quad E_{ii} = (\beta_{ii} + S_i^2 - S_i) / S_i, \text{ and}$$

$$(15) \quad E_{ij} = (\beta_{ij} + S_i S_j) / S_i \quad (i \neq j).$$

U.S. Food-Manufacturing Application: Data

To investigate the demand for labor, capital, and energy in the U.S. food-manufacturing industry, data are required on the unit price and total cost of each factor as well as on the value and quantity of industrial production. Most of the data are compiled from the *Annual Survey of Manufactures* (ASM), SIC code 20 (Food and Kindred Products) for 1971–86. Prior to this period, data for each sector of food manufacturing experienced serious definitional changes. The aggregate food sector includes the following nine categories: (1) meat products, (2) dairy products, (3) preserved fruits and vegetables, (4) grain mill products, (5) bakery products, (6) sugar and confectionery products, (7) fats and oils, (8) beverages, and (9) miscellaneous food and kindred products.

The value of food-manufacturing production is defined as the value of shipments adjusted for inventory changes of finished products. The aggregate price of the food-manufacturing product is defined as the Laspeyres price index for the producer price indexes of each food category (compiled from the *Producer Price Index*, PPI), using 1982 shipment values as weights. The aggregate price index is used to deflate the value of production to obtain an approximate quantity measure.

The cost of labor is the total wage payments for production workers, while unit labor price refers to the average wage payments per production worker per hour. The cost of energy is the total cost of purchased fuels and electric energy. The producer price index for fuel and power is used to represent the price of energy.

Implicit price deflators for gross fixed non-residential capital investment for structures and producers' durable equipment, compiled from

the *Survey of Current Business* (SCB), are used as an approximate price index for capital services. The cost of capital services for equipment and structures is the sum of depreciation, rental payments, and interest on the food-manufacturing net assets. Data for depreciation charges, rental payments, and the gross book value of depreciable assets are available in ASM. Because depreciation charges for 1972–76 and 1986 are not reported, this study projects them on the basis of a log-linear regression by fitting the depreciation charges (*D*) as a function of beginning-of-year structure and equipment assets (*K*) for 1977–85:

$$\ln D = -3.3195 + 1.0629 \ln K \\ (0.0292) \\ R^2 = 0.99 \\ D-W = 2.60.$$

Finally, the interest on net assets is calculated by multiplying the value of beginning-of-year assets by the Moody's Corporate Industrial Bond Rate from SCB.

Empirical Results

The cost structure of the U.S. food-manufacturing industry is characterized by estimating three cost share equations as in equation (10) for labor, capital, and energy. Because the cost shares of the three equations always sum to unity, the sum of the disturbances across these equations is zero at each observation. This implies that the covariance matrix of residuals is singular, and one of the cost share equations should be dropped from direct estimation. The remaining equations are then estimated simultaneously by applying Zeller's seemingly unrelated regression method, while the parametric constraints are incorporated.

As indicated in Berndt and Savin (p. 942), if one uses the covariance matrix of residuals obtained from applying ordinary least squares to unrestricted equations as prior information for estimation, the estimates of the parameters are the same regardless of which cost share equation is deleted. This study uses this approach, and the invariant property of estimates has been verified.

A time-trend variable as a proxy for technical progress was initially included in the model. However, because of multicollinearity, the estimated coefficients for the time trend were not statistically significant. Consequently, the time-trend variable was excluded from the model. In

addition, the relatively short sample period (15 years) precluded testing for structure changes in the food-manufacturing industry.

The estimation results are reported in table 1. Most of the estimated parameters (13 out of 15) are statistically significant at the 5% level. The percentages of root-mean-square error to sample mean are about 7% or less for each equation. These estimated parameters and their covariance matrix of errors together with the observed factor cost shares at sample means are the basic information for computing the estimates of factor demand relationships contained in tables 2 and 3. The factor cost shares used in the calculation are labor (0.5376), capital (0.3393), and energy (0.1231).

Based on equations (14) and (15), the (constant-output) price elasticities of factor demands

are obtained and shown in table 2. The results suggest that the demand for capital is quite elastic with an elasticity of 2.08, while the demand elasticities of labor and energy are 1.05 and 0.49, respectively. The high elasticity of demand for capital probably reflects the industry's high capital-intensive technologies. As indicated in Connor and others (p. 39), food processing is more capital intensive than such heavy industries as machinery and transportation equipment; only four major industry groups (paper, chemicals, petroleum, and primary metals) are more capital intensive than food manufacturing. The cross-price elasticities show a strong substitution relationship between capital and labor with elasticities 1.78 for the demand of capital and 1.12 for the demand of labor. Capital and energy are substitutable, and labor and energy are comple-

Table 1. Estimated Parameters of Cost Function

Equation	Constant	Factor Price of			Output Quantity	C.V. ^a
		Labor	Capital	Energy		
Share of:						
Labor	-0.77999* (0.92447) ^b	-0.31504 (0.05781)	0.42180 (0.06281)	-0.10677 (0.01201)	0.03897* (0.06553)	2.48
Capital	4.18320 (1.13008)	0.42180 (0.06281)	-0.48148 (0.07087)	0.05968 (0.01533)	-0.22115 (0.08114)	5.15
Energy	-2.40319 (0.35868)	-0.10677 (0.01201)	0.05968 (0.01533)	0.04708 (0.00608)	0.18218 (0.02732)	6.81

^a C.V. is the percentage of root-mean-square error to sample mean. All estimates except for those marked with (*) are significant at the 5% level.

^b Standard errors are in parentheses.

Table 2. Elasticities of Conditional Factor Demands and Allen Partial Elasticity of Substitution (AES)

Measure	Factor Price of		
	Labor	Capital	Energy
Elasticity			
Labor	-1.04849 (0.10754)	1.12401 (0.11685)	-0.07551 (0.02235)
Capital	1.78056 (0.18510)	-2.07953 (0.20886)	0.29897 (0.04519)
Energy	-0.32976 (0.09759)	0.82417 (0.12457)	-0.49441 (0.04940)
AES			
Labor	-1.95048 (0.20006)	3.31231 (0.34434)	-0.61344 (0.18155)
Capital		-6.12813 (0.61547)	2.42872 (0.36708)
Energy	(Symmetry)		-4.01639 (0.40133)

Note: Elasticities and their standard errors (in parentheses) are computed on the basis of equations from (12) to (15) at sample mean of cost shares. All estimates are significant at the 5% level.

mentary; however, the cross elasticities are relatively small.

The interdependencies among labor, capital, and energy are further confirmed by the Allen elasticities of substitution shown in the lower half of table 2. The elasticities signify substitution if the sign is positive and complementarity if the sign is negative. The substitution relationship between labor and capital is supported by previous studies of manufacturing industries. For example, Berndt and Christensen found that the Allen elasticities of equipment-labor and structures-labor ranged from 1.22 to 1.79 in U.S. manufacturing for 1929–68. Fuss found that the Allen elasticity between capital and labor was about 0.8 in Canadian manufacturing for 1961–71. In another study, Berndt and Wood found that the Allen elasticity between capital and labor was about 1.01 in U.S. manufacturing for 1947–71.

The Morishima elasticities of substitution calculated on the basis of equation (11) are compiled in table 3. The entries in the off-diagonal of the table reflect the adjustments of relative factors in response to a change in the ratio of relative factor prices. Their signs are all positive, implying that any pair of factors is substitutable with each other. In particular, the elasticities of factor ratios, labor-capital and capital-labor, are large, respectively, 2.83 and 3.20. The substitution between labor and energy is inconsistent with the Allen elasticity measure mainly because of different definitions; a Morishima

elasticity is related to the adjustment of two factors, while a partial adjustment of one factor is allowed in an Allen elasticity.

Both *MES* and *AES* indicate a strong substitution relationship between capital and labor. In fact, there is a trend of more intensive use of capital but less of labor in past years especially in light of the steady increase in the price ratio between labor and capital since 1982. The capital input index (1982 = 100) increased from 67.50 in 1972 to a peak of 110.15 in 1984 and then slightly decreased thereafter. On the other hand, the labor input index declined from 106.57 in 1972 to 95.02 in 1986.

The interrelatedness of factor demands is also shown in the variation of industrial cost structure in response to a change in factor prices. Based on equation (8), the elasticities of each pairwise factor cost share with respect to their factor prices are shown in the lower half of table 3. As indicated before, these results are closely related to the magnitude of Morishima elasticities; the relative cost share decreases if the Morishima elasticity is greater than one and increases if it is less than one. For example, the -1.83 elasticity of labor-capital indicates a significant reduction in the cost share of labor to capital in response to relatively higher wages than capital price. On the other hand, the 0.58 elasticity of energy-labor indicates that a marginal increase of energy price would cause an increase of the cost share of energy relative to labor.

Table 3. Morishima Elasticities of Substitution (*MES*) and Effects of Factor Price on Cost Shares

Measure	Factor Price of		
	Labor	Capital	Energy
<i>MES</i>			
Labor		2.82905 (0.29153)	0.71874 (0.11896)
Capital	3.20354 (0.32427)		2.90370 (0.30093)
Energy	0.41890 (0.04612)	0.79338 (0.08684)	
Cost share			
Labor		-1.82905 (0.29153)	0.28126 (0.11896)
Capital	-2.20354 (0.32427)		-1.90370 (0.30093)
Energy	0.58110 (0.04612)	0.20662 (0.08684)	

Note: Elasticities and standard errors (in parentheses) are computed on the basis of equations (8) and (11) at sample mean of cost shares. All estimates are significant at the 5% level.

Concluding Comments

This study analyzes the demand for labor, capital, and energy and their interdependent relationships in the U.S. food-manufacturing industry. The results show that the demand for capital services is more highly elastic than for labor and energy. Thus, any policy measures to reduce the price of capital services, such as investment tax credits and lower interest rates, would significantly increase the demand for capital. The demand elasticities of labor, capital, and energy in response to energy price changes are relatively low. They indicate that the relatively large changes in the prices of fuel and energy experienced over the sample period did not cause much adjustment in factor utilization.

The Morishima elasticity is, in general, better than the Allen elasticity for representing the factor substitution relationship because of its capability to explicitly explain the adjustment of factor combinations in response to relative price changes. The estimated Morishima elasticities indicate that labor, capital, and energy are substitutable especially between labor and capital. This is evidenced by the recent trends in the food-manufacturing industry to substitute computers and automated machines for human operations in light of the steady increase in the labor to capital price ratio.

[Received March 1990; final revision received November 1990.]

References

- Allen, R. G. D. *Mathematical Analysis for Economists*. London: Macmillan & Co., 1938.
- Berndt, E. R., and L. R. Christensen. "The Translog Function and the Substitution of Equipment, Structures, and Labour in U.S. Manufacturing 1929-68." *J. Econometrics* 1(1973):81-114.
- Berndt, E. R., and N. E. Savin. "Estimation and Hypothesis Testing in Singular Equation Systems with Autoregressive Disturbances." *Econometrica* 43(1975): 937-57.
- Berndt, E. R., and D. O. Wood. "Technology, Prices, and the Derived Demand for Energy." *Rev. Econ. and Statist.* 57(1975):259-68.
- Binswanger, H. P. "A Cost Function Approach to the Measurement of Elasticities of Factor Demand and Elasticities of Substitution." *Amer. J. Agr. Econ.* 56(1974):377-86.
- Blackorby, C., and R. R. Russell. "Will the Real Elasticity of Substitution Please Stand Up? (A Comparison of the Allen/Uzawa and Morishima Elasticities)." *Amer. Econ. Rev.* 79(1989):882-88.
- Connor, J. M., R. T. Rogers, B. W. Marion, and W. F. Mueller. *The Food Manufacturing Industries: Structure, Strategies, Performance, and Policies*. Lexington MA: Lexington Books, 1985.
- Fuss, M. A. "The Demand for Energy in Canadian Manufacturing: An Example of the Estimation of Production Structures with Many Inputs." *J. Econometrics* 5(1977):89-116.
- Hicks, J. R. *Theory of Wages*. London: Macmillan & Co., 1932.
- Lopez, R. E. "The Structure of Production and the Derived Demand for Inputs in Canadian Agriculture." *Amer. J. Agr. Econ.* 62(1980):38-45.
- Ray, S. C. "A Translog Cost Function Analysis of U.S. Agriculture, 1939-77." *Amer. J. Agr. Econ.* 64(1982):490-98.
- Shoemaker, R. *Effects of Changes in U.S. Agricultural Production on Demand for Farm Inputs*. Washington DC: U.S. Department of Agriculture, Econ. Res. Serv. Tech. Bull. No. 1722, 1986.
- U.S. Department of Commerce, Bureau of the Census. *Annual Survey of Manufactures*. Washington DC, various issues.
- U.S. Department of Commerce, Bureau of Economic Analysis. *Survey of Current Business*. Washington DC, various issues.
- U.S. Department of Labor, Bureau of Labor Statistics. *Producer Price Index*. data on tape.

Government Intervention in Imperfectly Competitive Agricultural Input Markets

Steve McCorriston and Ian M. Sheldon

The frequent demands for protection by the fertilizer industry in many developed countries may have important implications for the agricultural sector. This paper provides a theoretical justification for government intervention in the fertilizer market in the form of an import tariff against foreign competitors. However, a superior policy would be for the government to counter the competitive distortion that exists in the market. An even better outcome is attained by using both policies simultaneously. Moreover, since the fertilizer industry undergoes continuous structural change, the extent to which optimal policies should vary is also investigated.

Key words: fertilizer market, government intervention, imperfect competition, structural change.

In many developed countries, the domestic fertilizer industry often demands and receives government support. The bases for these requests include aid to improve the competitiveness of domestic manufacturers, to protect the industry as it undergoes rationalization, and to counteract the aggressive export policies of foreign competitors. An example of the former is the U.S. National Gas Policy Act in use over the early 1980s which lowered the manufacturing cost of ammonia (the main feedstock of most manufactured fertilizers) by an estimated 50% (European Commission). In the case of the fertilizer industry undergoing rationalization, the Japanese government imposed a number of measures to limit or exclude imports while a restructuring of ammonia, urea, and phosphoric acid facilities were being implemented (Lancaster). The latter case has been typified by the complaints in recent years by the European fertilizer industry to the EC Commission and national governments of unfair competition from Eastern bloc exporters. Some estimates report that exports from the Eastern bloc were selling at prices 25% below West European cost levels (Agra Europe).

Despite the implications of government intervention in the fertilizer industry for farmers, agricultural economists have largely paid little attention to these issues. The aim of this paper, therefore, is to explore whether there are any theoretical justifications for government intervention in agricultural input markets. The reference point for this analysis is the recent developments in the international trade literature which suggest that there may be a normative justification for government intervention in the presence of imperfect competition.

These theoretical developments are applied to the fertilizer industry using the UK market as a case study, and values for the optimal degree of protection (tariffs) are derived. However, two further issues are explored. First, the welfare effects resulting from trade policy intervention are compared with alternative policies. Specifically, the welfare effects of antitrust legislation (as proxied by a production subsidy) which corrects the competitive distortion arising in the domestic market are estimated. A final policy consideration is the simultaneous use of tariffs and antitrust legislation. Second, the theoretical justification for intervention rests on the imperfectly competitive nature of the fertilizer industry. However, market structure is not necessarily a constant variable. For example, in the case of the UK fertilizer market, firms have exited the

Steve McCorriston is a lecturer, University of Exeter, and Ian M. Sheldon is an associate professor, Ohio State University. Seniority of authorship is not assigned.

The authors appreciate the comments of the anonymous *Journal* reviewers on an earlier draft of this manuscript.

industry. Consequently, even if intervention is justified, the degree of intervention likely varies as market structure changes. Therefore, the response of optimal policies to changes in the structural characteristics of the fertilizer industry is also investigated.

The paper is organized as follows. The first section outlines the rationale for government intervention in imperfectly competitive markets. The structural characteristics of the fertilizer industry in developed countries are outlined in the second section. The theoretical framework, based on Dixit (1988), is presented in sections three and four. Section five calibrates the theoretical framework to the UK fertilizer market. The extent and effects of government intervention in the fertilizer market and how such government intervention should vary in response to changes in market structure and firms' behavior are then discussed.

Rent-Shifting Trade Policy

Recent developments in the international economics literature have explored the role for government policy when markets (home and foreign) are imperfectly competitive. The rent-shifting justification for an active trade policy is characterized by the work of Brander and Spence

(1984, 1985), Krugman, Dixit (1984), and Eaton and Grossman, among others. The basis for these arguments applied to an import market can be explained diagrammatically.

Two groups of firms (home and foreign) are assumed to share the domestic market. For simplicity, their costs are assumed equal and constant. With prices initially at P_1 , the domestic firms take OQ_1 of the market, while the foreign firms account for Q_1Q_2 (see fig. 1). With a tariff, the exporters' market share falls to Q_3Q_4 , while the domestic firms expand their market share to OQ_3 . As a result of the tariff, the price has now risen to P_2 and a terms of trade effect is denoted by $P_2 - t$. Domestic producers gain following an increase in profits ($A + D$), consumers lose ($-A - B - C$) and government revenue increases ($B + E + G$). The net welfare effect is $E + G + D - C$. The essence of these strategic trade policy arguments is that, despite the potentially significant losses to consumers, the increase in the domestic firms' profits coupled with the government's receipt of tariff revenue is sufficient to outweigh the consumer surplus loss.¹ Of course, this implies that the government applies equal weight to the three interest groups; if, however, the government applied a greater weight, say, to consumers' interests, then the simple rent-shifting argument for trade policy may become less desirable.

While the tariff counteracts the trade distortion, the domestic distortion of too little output persists. Consequently, if the government enforced antitrust legislation that ensured prices equalled costs, consumer surplus would increase by $I + D + E + F + H$ and the domestic firms would lose profits equal to $-I$, giving a net change in welfare of $D + E + F + H$. Clearly, if $F + H$ is larger than $G - C$, then antitrust legislation is a superior means of increasing national welfare compared to the optimal tariff. However, a first-best policy may be a combination of the tariff and antitrust legislation that counters both distortions simultaneously.

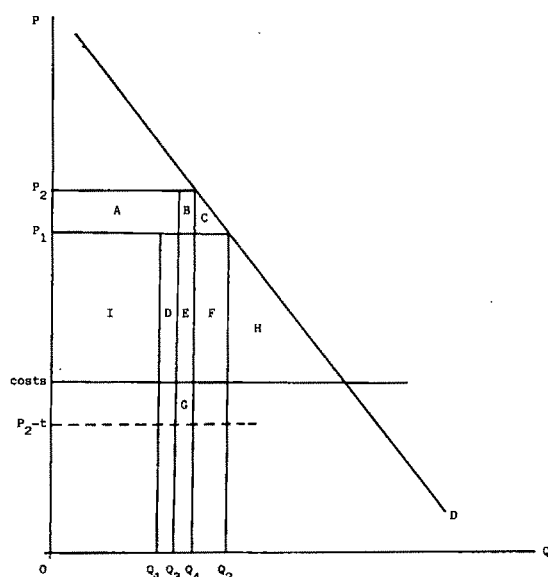


Figure 1. Fertilizer market analysis

¹ The outcome of this game is sensitive to the strategic variable because optimal policies may depend upon whether firms play Cournot (quantity strategies) or Bertrand (price strategies). Eaton and Grossman have shown that, for an exporter, Cournot conjectures would imply an optimal export subsidy while Bertrand conjectures would lead to an optimal export tax. However, in the context of an import market, tariffs are optimal for both Cournot and Bertrand conjectures, though, as Cheng has shown, the values for the optimal tariffs will differ.

Structural Characteristics of the Fertilizer Industry

While it is difficult to generalize, the fertilizer industry in many developed countries tends to be concentrated in nature. Access to raw materials such as natural gas, the capital-intensive nature of production, and economies of scale often create barriers to entry. Within the EC one or two dominant firms co-exist with a "competitive" fringe of smaller firms who blend fertilizers for local and regional markets (European Commission). For example, in the UK three firms now account for 80% of the market (UK Census of Production), with ICI taking just over 50%. Elsewhere in the EC, Ercros accounts for 60% of the Spanish market (*Financial Times*), BASF and Norsk Hydro (formerly VEBA) dominate the West German market, while Cdf Chemie and Norsk Hydro (formerly COFAZ) account for a substantial share of the French market (European Commission). The U.S. fertilizer industry is less concentrated (Leibensluft), with eight firms taking a 57% market share in 1982 (U.S. Census of Manufactures). Pressure from imports in the 1980s has forced some plant closures in the United States, which may have led to increased seller concentration (Daberkow).

In terms of market behavior and performance, there is some evidence that firms in developed countries have acted less than competitively. Shaw has traced price and product competition in the UK fertilizer industry in the postwar period, finding evidence for periods of oligopolistic accommodation as characterized by parallel pricing, interspersed by the occasional price war. Bayri, Rosaasen, and Furtan indicate that Canadian manufacturers have set prices in order to limit entry and maintain above-normal profits. However, the ability of firms to act less than competitively has been constrained by recent changes in the world fertilizer market. A shift in supply away from developed Western countries to developing and Eastern bloc countries has placed the former under greater competitive pressure. In particular, the EC market has been characterized by an inflow of low-priced imports (Agra Europe) resulting in increased demands by EC firms for protection from the predatory pricing tactics of non-EC firms.

Given these structural characteristics a *prima facie* case can be made for applying the theoretical arguments for government intervention when markets are imperfectly competitive to the fertilizer industry. However, as the structure of

the industry changes, so too should the extent of government intervention. For example, as the fertilizer industry becomes less competitive as a result of a change in its economic environment, one would expect a priori change in the level of the tariff and subsidy values for the policies to remain optimal. Typically, the factors affecting the fertilizer industries in developed countries in recent years have included changes in the number of fertilizer manufacturers, changes in costs (e.g., energy, wages, etc.) and increased competition from third-country exporters (dumping).

Derivation of Optimal Policies

The theoretical underpinnings for this paper are based upon Dixit (1988) and are fully explained in the appendix. The essential features of the model are as follows: the structure of the UK fertilizer market is divided into two, where subscript 1 refers to the dominant firms and subscript 2 refers to the blenders who act primarily as importers of urea (a substitute for ammonia); domestically produced fertilizers and imports are considered imperfect substitutes in agricultural production, these products being differentiated by nitrogen content per tonne; firm behavior is modeled as a Nash equilibrium in conjectural variations, although no explicit behavior (e.g., Cournot) is imposed on the model.

National welfare is the sum of farmers' producer surplus (Γ), the dominant firms' profits (Π_1) and government revenue/expenditure (t and s , where t is a tariff and s is a production subsidy, the proxy for antitrust policy), as given by

$$(1) \quad W = \Gamma + \Pi_1 + (tQ_2 - sQ_1).$$

The government's objective is to choose values for t and s that maximize national welfare. Essentially, the government has three policy choices: first, it can calculate a value for the optimum tariff without a production subsidy ($s = 0$); second, it can derive a value for the optimal production subsidy without a tariff ($t = 0$); or, finally, it can choose a joint policy combination. In the latter case, one policy will likely, in part, substitute for the role played by the other.

The optimal tariff without a subsidy is found by maximizing welfare with respect to t . Following Dixit, this gives an expression for the optimal tariff as

$$(2) \quad t = Q_2V_2 + (Q_1V_1)k/(b_1 + V_1).$$

Substituting for Q_1 and Q_2 from (A13), the explicit expression for the optimal tariff is

$$(3) \quad t = \frac{(a_1 - c_1)k(\beta_2 V_1 - \beta_1 V_2) + (a_2 - c_2)(\beta_1^2 V_2 - k^2 V_1)}{\beta_1^2(\beta_2 + V_2) - k^2(\beta_1 + V_1)},$$

where $\beta_i = (b_i + V_i)$.

The values for a_i , b_i , and k are taken from the inverse demand equations as shown in (A5) and (A4). The value for the optimal tariff is, therefore, dependent on the nature of competition in the market (V_i) and relative costs between the home industry and its competitors. With respect to firms' behavior, the higher the value for V_i , the higher the value of the optimal tariff. Intuitively, a high V_1 suggests that the home industry has market power. A higher V_2 suggests higher supernormal profits for the foreign industry. Thus, high values for V_i suggest a stronger role for tariffs in order to shift profits from the foreign industry to the home country.

The other determinant of the optimal tariff value is relative costs. In particular, since $(\beta_1^2 V_2 - k^2 V_1)$ is likely negative and $(\beta_2 V_1 - \beta_1 V_2)$ is positive, assuming that the values for a_i are similar, then differences between home and foreign costs are important. Thus, a home industry with a cost advantage strengthens the argument for a tariff, while a foreign industry with a cost advantage weakens the argument for a tariff. However, if costs are sufficiently similar and the two types of fertilizer are sufficiently imperfect substitutes for each other, then an optimal tariff will be justified.²

The optimal production subsidy is derived in a similar way to the tariff. Maximizing welfare with respect to s gives

$$(4) \quad s = Q_1 V_1 + (Q_2 V_2)k / (b_2 + V_2),$$

and substituting for Q_1 and Q_2 from (A12) gives

$$(5) \quad s = \frac{(a_1 - c_1)(\beta_2^2 V_1 - k^2 V_2) + (a_2 - c_2)k(\beta_1 V_2 - \beta_2 V_1)}{\beta_2^2(\beta_1 - V_1) + k^2(V_2 - \beta_2)},$$

where β_i is defined as above.

Again, government policy will vary with costs and the conjectural variations parameters. The lower the values for V_i , the lower the value for s ; i.e., because the market is tending towards perfect competition, there is less need for s to overcome the competitive distortion. The greater

the cost advantage of the home firm, the greater is the need for a production subsidy.

The joint optimum is derived by maximizing (1) with respect to t and s . Thus, the optimum tariff when s does not equal zero is

$$(6) \quad t = \frac{-(a_1 - c_1)kV_2 + (a_2 - c_2)b_1V_1}{(b_2 + 2V_2)b_1 - k^2}.$$

Similarly, the optimal subsidy when t does not equal zero is

$$(7) \quad s = \frac{(a_1 - c_1)V_1(b_2 + 2V_2) - (a_2 - c_2)kV_1}{(b_2 + 2V_2)b_1 - k^2}.$$

The determinants of the joint optimum tariffs and subsidies follow the intuition outlined above.

Adjustment of Policies to Structural Change

As indicated in section two, the fertilizer industry in developed countries is subject to frequent pressures. Thus, given that there is a normative justification for government intervention (in one form or another), it is worthwhile to consider how these policies should vary when the structural characteristics, i.e., c_i and V_i , of the fertilizer industry change. In particular, the effects of firm exit from the fertilizer industry are emphasized in this section.

Exit of firms from the market, say through take-over by a competitor, has been a recurring characteristic of structural change in the Euro-

pean fertilizer industry.³ In the UK market, the number of dominant firms has recently fallen from four to three as one firm (UKF) swapped its production facilities with Kemira in exchange for capacity in the Dutch market. Exit,

² Cheng has analyzed the possibility of corner solutions when costs are sufficiently differentiated and products are homogenous.

³ Particularly notable are the activities of the Norwegian-based Norsk-Hydro, which has acquired production capacity in most countries in Western Europe (McCorriston and Sheldon).

which will increase seller concentration, can be captured by a fall in n_1 . Hence, given $V_1 = -b_1/n_1$, the fall in the number of firms n_1 will lead to an increase in V_1 . Because the change in V_1 will be reflected in a change in p_1 , the change in the optimal tariff is obtained by deriving dt/dV_1 and dp_1/dV_1 to give

$$(8) \quad \left. \frac{dt}{dp_1} \right|_{s=0} = \frac{k\Delta'[\beta_2(\Delta' - 2\beta_2V_1) + V_2\Delta' - (2V_2\beta_1 - k^2)]}{2(\beta_2b_1 - k^2)[\beta_1^2(\beta_2 + V_2) - k^2(\beta_1 + V_1)]},$$

where $\Delta' = (\beta_1\beta_2 - k^2)$, and

$$(9) \quad \left. \frac{dt}{dp_1} \right|_{s \neq 0} = \frac{b_1(a_2 - c_2)\Delta'}{[(b_2 + 2V_2)b_1 - k^2]2Q_1(\beta_2b_1 - k^2)},$$

where Δ' is given as above and Q_1 is the initial level of output.

By a similar procedure, the change in the optimal subsidy will be given as

$$(10) \quad \left. \frac{ds}{dp_1} \right|_{t=0} = \frac{\Delta'(2b_2 + V_2)}{2(\beta_2b_1 - k^2)\beta_2^2b_1 - k^2b_2}, \text{ and}$$

$$(11) \quad \left. \frac{ds}{dp_1} \right|_{t \neq 0} = \frac{[(a_1 - c_1)(b_2 + 2V_2) - k(a_2 - c_2)]\Delta'}{[(b_2 + 2V_2)b_1 - k^2]2Q_1(\beta_2b_1 - k^2)},$$

where Δ' and Q_1 are as defined above.

Because the rise in V_1 is indicative of greater concentration in the fertilizer industry, the value for both optimal policies is expected to increase.

Besides exit of firms, other changes in market structure will also affect the values of optimal policies. These will include changes in domestic and foreign firms' costs and the behavior of foreign firms. The direction of policy adjustments is summarized in table 1. The structural changes shown have been typical of those that have affected the European fertilizer industry in recent years; domestic costs have fallen because of increased technical efficiency; exporters' costs have

fallen as a result of government subsidies (European Commission); lower oil prices in the mid-1980s reduced both the costs of domestic and foreign firms; and accusations of dumping by exporters characterizes changes in foreign competitors' behavior.

Model Calibration

In order to quantify the optimal policies outlined above, estimates of the parameters in the demand system (A3) and (A4) are needed. However, equations (A5) and (A6) indicate five unknowns (A_1 , A_2 , B_1 , B_2 , and K) but only two relations between them. Consequently, three further relations are required to solve the system. Following Dixit (1987), expressions for the elasticities of demand and substitution are derived and set equal to empirically observed values.

Because the products of the dominant firms and of the blenders are treated as imperfect substitutes, the price elasticity of demand is interpreted as the effect of an equiproportionate rise in the price of the two products on total fertilizer expenditure Q . Therefore letting $p_1 = P_1^0P$ and $p_2 = P_2^0P$, where P_1^0 and P_2^0 are initial prices and P is the proportional change factor, the aggregate expenditure for fertilizers can be written as

$$(12) \quad Q = P_1^0Q_1 + P_2^0Q_2.$$

Given that p_1 and p_2 are the initial prices in the calibration, and substituting equations (A5) and (A6) into (28), the aggregate expenditure index can be written as

$$(13) \quad Q = p_1A_1 + p_2A_2 - (B_1p_1^2 + B_2p_2^2 - 2Kp_1p_2)P.$$

Table 1. Adjustment of Policies to Structural Change

	Tariff ($s = 0$)	Subsidy ($t = 0$)	Tariff ($s \neq 0$)	Subsidy ($t \neq 0$)
Domestic firm ceases production (exit)	+	+	+	+
Domestic costs fall	+	+	+	+
Foreign firms' costs are subsidized	+	-	+	-
Exogenous fall in energy costs	+	+	+	+
Dumping by foreign exporters	-	-	-	-

Note: The derivation of these policy adjustments are available from the authors on request.

The total market elasticity of demand for fertilizer, ϵ , is then defined and evaluated at the initial point where the proportional change factor P equals 1. By differentiating (13) with respect to P and multiplying by P/Q , the elasticity is given as

$$(14) \quad \epsilon = - \frac{B_1 p_1^2 + B_2 p_2^2 - 2K p_1 p_2}{Q}$$

Expression (14) is then set equal to the observed value of ϵ .

The elasticity of substitution would normally be defined as

$$(15) \quad \sigma = d \log(Q_1/Q_2) / d \log(p_1/p_2),$$

which gives a fourth relation between the parameters when set equal to an empirically observed value for σ . However, equations (A5) and (A6) in general define the ratio Q_1/Q_2 as a function of the vector (p_1, p_2) and not in terms of the ratio p_1/p_2 . In order for Q_1/Q_2 to be a function of p_1/p_2 , at least locally, the parameters must satisfy the following final relation:

$$(16) \quad p_1(A_1 K + A_2 B_1) = p_2(A_2 K + A_1 B_2),$$

which implies homotheticity of the production function. Given the definition of σ in (15) and using equations (A5), (A6), and (16), the final expression for the elasticity of substitution can be derived as

$$(17) \quad \sigma = \frac{\frac{p_1}{p_2} (B_1 B_2 - K^2)}{\left(B \frac{p_1}{p_2} - K \right) \left(B_2 - K \frac{p_1}{p_2} \right)}$$

The simultaneous equation system using (A5), (A6), (14), (16), and (17) is solved to obtain estimates of A_1 , A_2 , B_1 , B_2 , and K where prices and quantities and estimates of the elasticity of demand and substitution are taken from outside sources. This solution reflects the process of model calibration such that the parameters are consistent with equilibrium in any given period. Values for a_1 , a_2 , b_1 , b_2 , and k can now be derived. Because of the limited availability of data, the model was calibrated for the UK fertilizer market for the year 1985: p_1 (£126) and p_2 (£120) are the average selling prices of the dominant firms and blenders, respectively, over the year 1986 based on reported prices in the UK farming press; Q_1 and Q_2 (1,268 and 317 thousand tonnes, respectively) are derived from Fertilizer Manufacturers Association data and other farm-

ing and trade sources. The elasticity of demand ϵ (assumed to be -0.65) is based on an estimate made by Metcalf and Cowling, although more recent estimates by Burrell suggest a similar value. No estimate of σ is available for the UK, so a value of 2.00 was assumed; the 2.00 value compares with an Australian estimate made by Higgs of 1.7. The cost estimates are £100 for c_1 , for the dominant firms based upon reported cost levels in Challinor and the UK farming press, and c_2 at £110 representing blenders' operating costs, which are assumed to be £10 below the selling price p_2 . The parameter estimates for the model are shown in table 2.

Welfare Effects of Government Intervention

To recap the discussion in the first section, if markets (home and foreign) are characterized by imperfect competition an activist trade policy may increase national welfare. However, production subsidies may represent a superior policy instrument, although a combination of tariffs and subsidies may be the most desirable policy. Using the calibrated parameters and cost information, values for the optimal tariff and subsidy policies were derived from (3), (5), (6), and (7), respectively.

In line with the theoretical predictions the optimal tariff has a value of £18.21 per tonne of fertilizer. However, the optimal production subsidy is £31.79 per tonne. When the tariff and subsidy are chosen simultaneously, the policies in part substitute for each other. The optimal tariff is now £2.71 per tonne, while the optimal subsidy is reduced to £30.38 per tonne. Solving for the new prices and quantities following the imposition of policy using (A3), (A4), and (A11), and (A12), the effects on national welfare can be estimated using (1). The results are presented in table 3 and are compared with the prepolicy welfare.

As the new trade theory predicts, the optimal tariff should increase overall welfare. However,

Table 2. Demand Parameters

Aggregate Demand Functions		Inverse Demand Functions	
A_1	2,092,200	a_1	320
A_2	523,050	a_2	305
B_1	7799	b_1	(10^{-4}) 1.38
B_2	3104	b_2	(10^{-4}) 3.47
K	1321	k	(10^{-5}) 5.88

Table 3. Welfare Effects of Optimal Tariff and Subsidy Policies (£m)

	Farmers' Welfare	Dominant Firms' Profits	Government Revenue/Expenditure	Overall Welfare
Pre-policy welfare	152.36	32.97		185.33
Welfare with optimal tariff ($t = £18.21$)	147.23	34.06	4.86	186.15
Welfare with optimal subsidy ($s = £31.79$)	183.49	48.93	-43.90	188.52
Welfare with joint optimal tariff and subsidies ($t = £2.71, s = £30.38$)	181.47	50.08	-43.03	188.53

the overall welfare gains are small (0.44% relative to the prepolicy case) and mainly involve a redistribution of welfare from farmers' to the dominant firms and the government. However, as expected, the optimal production subsidy leads to a higher level of economic welfare by countering the competitive distortion in the fertilizer market.⁴ The net welfare gains, in this case, although still low (1.77% relative to the prepolicy case) are larger than those resulting from the optimal tariff. The joint use of the tariff and subsidy policies is only marginally superior to using the subsidy alone.

Despite the small net welfare changes, the optimal production subsidy leads to a considerable redistribution of welfare between farmers, the dominant fertilizer manufacturers, and the government. In particular, because the optimal subsidy counters the oligopolistic distortion, farmers' welfare increases as a result of the lower prices for fertilizers. Essentially, farmers' producer surplus increases by around £36m in 1985 or by 25% relative to prepolicy welfare. The dominant firms' profits also increase substantially (by 43%), while the government must pay for the subsidy. Although the welfare gain arising from the joint policies is only marginally better than using the subsidy alone, the main difference between these two cases involves the distribution of income. Farmers' gains are lower and the dominant firms' profits greater in the joint optimum case relative to using the subsidy alone. The sensitivity of the results to the data used to calibrate the model is considered below.

Elasticity Data

Model calibration relies upon estimates of the elasticities of demand and substitution. In order to analyze how sensitive the results are to changes

in these elasticities, different values for ϵ and σ were chosen. The corresponding optimal policies and welfare outcomes are presented in table 4.

The central column shows the base results derived above. As the elasticities increase (in absolute terms), the values for the optimal tariffs and subsidies increase, except for the value of the tariff in the joint optimum. However, the ranking of the policies is not sensitive to the elasticity data. Specifically, tariffs lead to a very small increase in national welfare, although more substantial increases accrue from the subsidy policy. Again, the joint optimum case is better, but only marginally so, relative to the subsidy alone.

Economies of Scale

An important feature of the fertilizer industry in developed countries is scale economies. In the model's derivation, however, this feature largely was ignored because of the problem of allocating capital costs intertemporally in the cost function. Nevertheless, it is still interesting to enquire whether the results are sensitive to firms reaping economies of scale.

An external estimate of economies of scale is used to determine how far costs may fall as output increases as a result of the various policies. Unfortunately Pratten (the main source of information on UK and European economies of scale) provides no estimate for the fertilizer industry. Instead, it is assumed that the extent of scale economies in the fertilizer industry is similar to that in the petrochemical industry which implies a cost elasticity of 20% (Pratten). With this proxy estimate, the likely changes in firms' profits (the main source of benefit arising from the scale economies) was estimated for the various forms of government intervention. The results are presented in table 5.

The results from table 5 suggest that an allowance for (extensive) economies of scale would

⁴ In practice, the welfare effects of antitrust legislation would not involve the transfer of a subsidy payment from government to producers.

Table 4. Sensitivity Analysis: Changes in Elasticity Data

		$\epsilon = -0.5$ $\sigma = 1.75$	$\epsilon = -0.65$ $\sigma = 2.00$	$\epsilon = -0.80$ $\sigma = 2.75$
Pre-policy welfare	(£m)	230.78	185.33	156.60
Optimal tariff	(£)	16.88	18.21	26.60
Welfare with tariff	(£m)	231.10	186.15	157.44
Optimal subsidy	(£)	30.36	31.79	33.83
Welfare with subsidy	(£m)	232.90	188.52	160.75
Joint optimum:				
Tariff	(£)	3.70	2.71	0.22
Subsidy	(£)	29.18	30.38	31.96
Welfare with joint optimum	(£m)	232.95	188.53	160.75

Table 5. Sensitivity Analysis: Economies of Scale (Cost Elasticity = 0.2)

	Domestic Firms' Profits ^a (£m)	
	No Economies of Scale	Economies of Scale
Optimal tariff	34.06	34.09
Optimal subsidy	48.93	51.01
Joint optimum	50.98	50.28

^a Original profits were £32.97m.

not significantly affect the welfare calculations. The greatest change in the firms' profits occurs for the optimal subsidy; profits increase by 54.7% when costs decrease compared with 48.4% when costs remain constant. This 54.7% increase in profits translates into a 2.8% increase in welfare (relative to the prepolicy levels) compared with a 1.72% increase in welfare in the constant costs case.

Cost Estimates

Finally, because the level of costs is important in determining the level of the policy variable, it is important to consider how sensitive the policy values are to changes in costs. The focus is on changes in the tariff values (in the absence of the subsidy) in response to changes in the costs of the domestic firms.

Two cases are compared with the base outcome. The first is where the costs are lower than the original case (£90 per tonne as opposed to £100 per tonne), and the second is where the domestic firms' costs are equal to those of the importers (i.e., £110 per tonne). The corresponding values for the tariff levels are shown in table 6. The results indicate that the value for

Table 6. Sensitivity Analysis: Effect of Changes in Domestic Firms Costs on Tariff Levels (£)

	$c_1 = 90$	$c_1 = 100$	$c_1 = 110$
Optimal tariff ($s = 0$)	21.11	18.21	15.02

the optimal tariff increases (decreases) as costs rise (fall). Similar results occur for the optimal subsidy. These results imply that higher costs are associated with more competitive conduct (lower V_i) and, therefore, less need for government intervention.

In sum, government intervention in the fertilizer industry can be normatively justified. The sole use of tariffs is an inferior policy relative to antitrust policy (optimal subsidies), although the best outcome (though only marginally so) is achieved by the joint use of tariffs and subsidies. While the overall increase in welfare (assuming equal weights between the various groups) is generally small, the main effect of these policies is on the distribution of income, particularly between farmers and the domestic fertilizer manufacturers. Farmers would gain substantially from antitrust legislation, while their welfare would be significantly reduced from (optimal) trade policy intervention.

Optimal Policies and Changes in Market Structure

These optimal policies likely will vary as market structure changes. Therefore, following the earlier analysis of section 4, values for the optimal policy responses to various market scenarios are presented in table 7. The values in table 7 represent response by tariffs or subsidies

Table 7. Tariff and Subsidy Policy Changes as Market Structure Varies (per unit change)

	Tariff ($s = 0$)	Subsidy ($t = 0$)	Tariff ($s \neq 0$)	Subsidy ($t \neq 0$)
Domestic firm ceases production (exit)	1.28	0.121	0.2315	0.2592
Domestic costs fall	0.019	0.152	4.62 ⁻⁰⁶	0.1582
Foreign firms' costs are subsidized	0.074	-0.009	0.0533	-0.0227
Exogenous fall in energy costs	0.143	0.093	0.0184	0.1355
Dumping by foreign exporters	-0.504	-0.121	-0.4956	-0.0269

to a unit change in, say, domestic costs. All policy responses fulfill the a priori expectations outlined above. For example, if foreign firms' costs are subsidized by their government, an increase in tariff would countervail this subsidy. However, due to the assumption of a constant price-cost mark-up for the blenders, the lower costs are reflected in lower market prices. Hence, imports provide a procompetitive discipline on the fertilizer market. When foreign firms dump their fertilizer exports on the market, both policy values should fall as a result of the competitive effect on the market. The other scenarios imply that the tariff and subsidy values should increase because the market is less competitive. In the joint optimum case, changes in the tariff are lower relative to changes when the tariff is the sole policy. However, except for dumping by foreign exporters, changes in the subsidy levels were greater in the joint optimum case compared to the subsidy case alone.

These results imply that, given the normative case for government intervention in the fertilizer industry, the levels of government intervention should adjust to change in the fertilizer market; however, the extent of the adjustment depends upon the source of change in the fertilizer market.

Consider an example where the focus is on changes in the level of welfare (following a change in market structure) if the government did not change the policy value relative to the case of fine-tuning its policy. The example is a domestic fertilizer manufacturer that ceases production. This scenario is representative of the UK market when UKF, one of four dominant domestic firms, ceased operation in the UK in 1988. The welfare effects of the UK government adjusting and not adjusting its intervention policies are shown in table 8.

In the benchmark case the government is initially using optimal tariffs and/or subsidies. If, as firms exit, the government did not change the policy values, economic welfare would be reduced in all cases. Thus, in the case of using the tariff alone, with no change in the optimal tariff, welfare would fall by 5.3%. The corresponding changes in welfare in the optimal subsidy and joint optimum cases are -0.75% and -2.7%, respectively. However, if the government did change the policy values by the extent recorded in table 7, then the optimal tariffs and subsidies would rise in all cases (given that the fertilizer market is now less competitive). The effects of these changes are either that the welfare losses are minimized (as in the case of us-

Table 8. Effects of Firm Exit

	Welfare £m	Percent Change Relative to No-Exit Case
<i>Tariff: $t = £18.21$</i>	186.15	
Exit: $t = £18.21$	176.23	-5.3
Exit: $t = £57.71$	183.54	-1.4
<i>Subsidy: $s = £31.79$</i>	188.52	
Exit: $s = £31.79$	187.11	-0.75
Exit: $s = £35.54$	188.57	0.03
<i>Joint optimum: $t = £2.71, s = £30.38$</i>	188.53	
Exit: $t = £2.71, s = £30.38$	183.51	-2.7
Exit: $t = £9.8, s = £38.40$	189.65	0.6

ing the tariff alone) or economic welfare is marginally increased (as shown in the optimal subsidy and joint optimum cases).

In sum, failure to change the optimal policy values in response to changes in market structure may lead to significant welfare losses. Thus, if government intervention in agricultural input markets is warranted, then the extent of intervention should respond to changes in the economic and policy environment.

Concluding Comments

Fertilizer manufacturers in developed countries often demand support from their home governments to improve the competitiveness of their industry or to protect it from external pressures. By applying the recent developments in the international economics literature, it has been shown that governments may use an optimal tariff to protect their domestic producers on the grounds that rents can be shifted from their international competitors. However, a stronger case exists for governments to use an optimal production subsidy (a proxy for antitrust policy) to counter the oligopolistic distortion in the fertilizer industry. The best policy is to use tariffs and subsidies together.

These theoretical results were validated by providing quantitative estimates for these policies for the UK fertilizer market. While the net welfare changes were small, both policies involved a substantial redistribution among farmers, the dominant firms, and the government. Farmers would experience substantial gains if antitrust policy promoted greater competition in the market, while considerable losses would occur under an optimal tariff. Using both policies simultaneously improves welfare only marginally more than using the subsidy alone, though the benefits to farmers are slightly less. Further, it was argued that governments should remain active following the initial imposition of the optimal policies. In particular, as the market becomes more (less) competitive, the optimal tariff and subsidy values should fall (increase).

The overall conclusion of this paper is that agricultural economists should pay more attention to the demands for government support that originate in the agricultural inputs sector. In particular, while there may be a normative justification for government protection against imports, higher levels of economic welfare can be attained if governments counter the competitive

distortion that is characteristic of many agricultural input markets.

[Received November 1989; final revision received October 1990.]

References

- Agra Europe. *East Europe Agriculture*. June 1987.
- Bayri, T., K. A. Rosaasen, and W. H. Furtan. "Limit Pricing in the Nitrogen Fertilizer Market: An Application to the Saskatchewan Market," abstract. *Amer. J. Agr. Econ.* 68(1986):1376.
- Brander, J. A., and B. J. Spencer. "Export Subsidies and International Market Share Rivalry." *J. Int. Econ.* 18(1985):83-100.
- . "Tariff Protection and Imperfect Competition." *Monopolistic Competition and International Trade*, ed. H. Kierzkowski. Oxford: Basil Blackwell, 1984.
- Burrell, A. "Demand for Fertilisers in the United Kingdom." *J. Agr. Econ.* 40(1988):1-20.
- Challinor, S. "The UK Fertiliser Industry: Situation and Outlook." *Agricultural Input Industries*, ed. J. Lingard. Newcastle: University of Newcastle, 1987.
- Cheng, L. K. "Assisting Domestic Industries Under International Oligopoly: The Relevance of the Nature of Competition to Optimal Policies." *Amer. Econ. Rev.* 78(1988):746-58.
- Daberkow, S. "Agricultural Input Industry Indicators in 1974-85." Washington DC: U.S. Department of Agriculture, Econ. Res. Serv. Agr. Info. Bull. No. 534, 1987.
- Dixit, A. "Anti-Dumping and Countervailing Duties Under Oligopoly." *Eur. Econ. Rev.* 32(1988):55-68.
- . "International Trade Policy for Oligopolistic Industries." *Econ. J. Suppl.* 94(1984):1-16.
- . "Optimal Trade and Industrial Policy for the U.S. Automobile Industry." *Empirical Methods in International Trade*, ed. R. Feenstra. Cambridge MA: MIT Press, 1987.
- Eaton, J., and G. M. Grossman. "Optimal Trade and Industrial Policy Under Oligopoly." *Quart. J. Econ.* 100(1986):383-406.
- European Commission. *The Community Fertiliser Industry*. Luxembourg, 1985.
- Fertiliser Manufacturers Association. *Fertiliser Review*, various issues.
- "EC Therapy Seems to Work." *Financial Times*, 19 Feb. 1990.
- Higgs, P. J. *Adaptation and Survival in Australian Agriculture*. Oxford: Oxford University Press, 1986.
- Krugman, P. R. "Import Protection as Export Promotion." *Monopolistic Competition in International Trade*, ed. H. Kierzkowski. Oxford: Basil Blackwell, 1984.
- Lancaster, J. "International Fertiliser Trade—Laissez-Faire or Protection?" *Fertiliser Review 1989*. London: Fertiliser Manufacturers Association, 1989.
- Leibenluft, R. F. "Competition in Farm Inputs: An Ex-

amination of Four Industries." Washington DC: *Federal Trade Commission*, Feb. 1981.

Metcalf, D., and K. Cowling. "Demand Functions for Fertiliser in the UK, 1948-65." *J. Agr. Econ.* 18(1967):375-86.

McCorriston, S., and I. M. Sheldon. "EC Integration and the Agricultural Supply Industries" *European Integration and Industry* ed. M. Macmillan, D. G. Mayes, and P. van Veen. Tilburg: Tilburg University Press, 1987.

Pratten, C. *A Survey of Economies of Scale*. Dep. Appl. Econ., University of Cambridge, 1988.

Shaw, R. W. "New Entry and the Competitive Process in the UK Fertiliser Industry." *Scottish J. Polit. Econ.* 27(1980):1-16.

U.K. *Census of Production*. London: Her Majesty's Stationary Office, 1988.

U.S. *Census of Manufactures*. Washington DC: U.S. Department of Commerce, Bureau of Census, 1982.

Appendix

Theoretical Framework

The theoretical framework for this paper is based upon Dixit (1988). The structure of the UK fertilizer market is divided into two, where subscript 1 refers to the dominant firms and subscript 2 refers to the blenders, who act as importers of urea (a substitute for ammonia). Because the blenders are treated as import agents, for simplicity they do not enter into the national welfare function. Initially, no entry/exit of firms occurs. Further, the dominant firms face constant average and marginal operating costs and blenders have a constant price-cost mark-up. Also, domestically produced fertilizers and imports are considered imperfect substitutes in production because these products have different nitrogen content per tonne.

Fertilizer market. Farmers' demand for fertilizers can be derived by maximizing the following profit function Γ for farmers with respect to Q_1 and Q_2 :

$$(A1) \quad \Gamma = f(Q_1, Q_2) - p_1 Q_1 - p_2 Q_2,$$

where the aggregate production function $f(Q_1, Q_2)$ for farmers is defined as

$$(A2) \quad f(Q_1, Q_2) = a_1 Q_1 + a_2 Q_2 - 1/2 (b_1 Q_1^2 + b_2 Q_2^2 + 2k Q_1 Q_2).$$

For simplicity, this production function is of quadratic form, no inputs other than the two forms of fertilizer Q_1 and Q_2 are considered, and farmers' output prices are normalized to one.

The farmers' profit function can be used to derive the respective inverse demand functions for the dominant firms and blenders, as given below:

$$(A3) \quad p_1 = a_1 - b_1 Q_1 - k Q_2, \text{ and}$$

$$(A4) \quad p_2 = a_2 - k Q_1 - b_2 Q_2,$$

where all parameters are positive, $(b_1 b_2 - k^2) > 0$ since

the products are imperfect substitutes, p_1 and p_2 are prices, and Q_1 and Q_2 are quantities. The corresponding demand functions are given as

$$(A5) \quad Q_1 = A_1 - B_1 p_1 + K p_2, \text{ and}$$

$$(A6) \quad Q_2 = A_2 + K p_1 - B_2 p_2,$$

where $B_1 B_2 - K^2 > 0$.

On the fertilizer production side, n_i firms are in each sector of the industry. Profits for a representative firm are given by

$$(A7) \quad \Pi_1 = (p_1 - c_1 + s) q_1 - f_1, \text{ and}$$

$$(A8) \quad \Pi_2 = (p_2 - c_2 - t) q_2 - f_2,$$

where prices and quantities are defined above, c_i is the marginal operating costs, and f_i is fixed costs.⁵ A production subsidy is given to the firm in the dominant sector of the market (the home firm), and t is a tariff on imports from foreign competitors.

The behavioral assumption is that firms' reactions to one another are treated as a Nash equilibrium with conjectural variations which arise from the firms' profit functions. Despite the problems with this approach, using conjectural variations in this model requires no explicit form of oligopolistic interdependence. Rather, the values of these parameters come from the data.

The conjectural variation parameter is derived from the first-order condition of the respective profits functions as given below:

$$(A9) \quad p_1 - c_1 + s + q_1 dp_1/dq_1 = 0, \text{ and}$$

$$(A10) \quad p_2 - c_2 - t + q_2 dp_2/dq_2 = 0,$$

where dp_i/dq_i is the conjectural variation parameter, i.e., the firm's expectation of how market price will vary with changes in output. Therefore, if a representative firm plays Cournot, it believes rival firms will not change output in response to a change in q_i ; hence, $dp_i/dq_i = -b_i$, the slope of the inverse demand function. If the market were perfectly competitive, a change in one firm's output would have no effect on market price, i.e., $dp_i/dq_i = 0$.

These first-order conditions for representative firms can be aggregated over n_i firms in each sector of the industry to give

$$(A11) \quad p_1 - c_1 + s + Q_1 V_1 = 0, \text{ and}$$

$$(A12) \quad p_2 - c_2 - t + Q_2 V_2 = 0,$$

where V_i is the aggregate conjectural variations parameter. Thus, for Cournot behavior, $V_i = -b_i/n_i$, and, as n_i increases, the more competitive the Cournot outcome becomes. In the limit $V_i = 0$, i.e., perfect competition.

Equilibrium prices and quantities in this fertilizer market model are obtained by combining (A3) and (A4) with (A11)

⁵ The profits functions in (A7) and (A8) include fixed costs. However, because data on this component of costs are not available, it is assumed that operating costs include fixed costs.

and (A12). Explicit solutions for prices and quantities are given below:

(A13)
$$\begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} = \frac{1}{\Delta'} \begin{bmatrix} b_2 + V_2 & -k \\ -k & b_1 + V_1 \end{bmatrix} \begin{bmatrix} a_1 - c_1 + s \\ a_2 - c_2 - t \end{bmatrix}$$

$\begin{bmatrix} a_1 - c_1 + s \\ a_2 - c_2 - t \end{bmatrix}$, and

(A14)
$$\begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} - \frac{1}{\Delta'} \begin{bmatrix} \Delta + b_1 V_2 & k V_1 \\ k V_2 & \Delta + b_2 V_1 \end{bmatrix} \begin{bmatrix} a_1 - c_1 + s \\ a_2 - c_2 - t \end{bmatrix}$$

where $\Delta = (b_1 b_2 - k^2)$ and $\Delta' = (b_1 + V_1)(b_2 + v_2) - k^2 = (\beta_1 \beta_2 - k^2)$.

Supply Impact of the Milk Diversion and Dairy Termination Programs

Bruce L. Dixon, Dwi Susanto, and Calvin R. Berry

A two-equation random coefficient regression model of commercial milk production is estimated using monthly observations from 1983 through June 1988 for the twenty-one major milk-producing states in the United States. Policy variables are entered into the model to represent the impact of the Milk Diversion Program (MDP) and the Dairy Termination Program (DTP). The MDP was primarily short term in impact and the DTP has been longer term. The DTP provision banning producers who exited under the program from producing for five years appears to have been ineffectual. The programs' effectiveness across states has varied considerably.

Key words: milk diversion and termination program effectiveness, random coefficients.

For over thirty-five years the U.S. dairy industry has been characterized by intensive government interventions. Under this regulation, the industry has provided adequate quantities of milk at relatively stable producer and consumer prices (Whipple, Powe, and Gray; Berry and Blakley). However, the situation changed during the early 1980s when the dairy industry began producing expanding surpluses. Milk production has been showing a steady tendency to rise.

The increase in milk production resulted in large government outlays through the dairy price support program or other incentives (Whipple, Powe, and Gray). A reduction in the price support level and a per hundredweight assessment were legislated in 1983. The purposes of these actions were to reduce milk production by lowering producer prices and to support promotion and product development.

Two major voluntary supply control programs, the Milk Diversion Program (MDP) and the Dairy Termination Program (DTP), were implemented in the mid-1980s. The MDP compensated dairy farmers who contracted to restrict their output by 5% to 30% of their historical base by paying them \$10 per hundredweight for the contracted reduction from their historical base. The MDP was in effect from 1 January

1984 through 31 March 1985. Dairy farmers subscribed to remove 7.5 billion pounds (Miller and Carman) of production on an annual basis.

The DTP required that producers whose bids were accepted sell or export all of their female dairy cattle. Moreover, the dairyman could have no interest in milk production for the ensuing five years, nor could his facilities be used for milk production in the five-year period. In addition to selling the entire herd for either slaughter or export, the farmer received a payment from the government. This payment was based on the bid per hundredweight (cwt.) of milk reflected in the farmer's offer in return for exiting the industry for at least five years. The bid averaged approximately \$15.00 per hundredweight of annual production for the U.S. (Matthews). These bids were accepted on a competitive basis nationwide and resulted in 11.3 billion pounds of annual capacity exiting the industry between April 1986 and August 1987 (ASCS, personal communication).

Both the MDP and DTP were intended to decrease the production of milk. Consequently, it is important to determine the short- and longer-term impacts of these programs on milk production. Because milk supply is likely to increase in the future, especially considering the introduction of bovine somatotropin (Kaiser and Tauer), it is important to measure the effectiveness of past programs since they are possible candidates for future programs. Moreover, because of the heterogeneity of the dairy industry across the United States, it is important to measure the diversity of the impacts by region.

Bruce L. Dixon and Calvin R. Berry are professors of agricultural economics, Department of Agricultural Economics and Rural Sociology, University of Arkansas; Dwi Susanto is a lecturer at the University of Lampung, Indonesia.

The computing assistance of Siew Goh and Diana Danforth is gratefully acknowledged, as are the helpful comments of the *Journal* referees.

To estimate the impact of the MDP and DTP, a two-equation model is specified and estimated using time-series and cross-sectional data. The first equation is yield per cow, and the second equation is cow numbers. Data are monthly for each of the twenty-one states listed by the U.S. Department of Agriculture (USDA) as the major milk-producing states in the United States. Using monthly data provides for a much more accurate measure of program effectiveness since the MDP lasted only fifteen months, and the exit period for the DTP lasted seventeen months. Hence, models using annual or quarterly data would have greater difficulty capturing the dynamic responses engendered by the programs. Both equations are specified as Swamy random coefficient regressions (RCR).

In what follows, an econometric model of commercial milk production is specified and estimated. The forecasting ability of the model is verified through a post-sample prediction experiment. The model is used to simulate milk production, first under the assumption that both the MDP and DTP were in effect, then second that only the MDP was implemented, and third, assuming only the DTP was implemented. A fourth simulation assumes neither were implemented and compares production with the predicted production when the various combinations of programs were implemented. Implications of these simulations are then considered.

Model Specification

Following prior studies (Wilson and Thompson; Chen, Courtney, and Schmitz; Kaiser, Streeter, and Liu) commercial milk production is modeled as the product of number of cows freshened (*CN*) and production per cow (*PPC*). Both equations are double logarithmic. The *CN* equation is a function of the ratio of the blend price of milk to the price of 16% protein dairy feed (*MFPR*), both variables lagged one month, variables representing the amount of base production withdrawn during the two programs, and a lagged dependent variable. Algebraically the *CN* equation is written as

$$(1) \quad \ln(CN)_{it} = \alpha_0 + \alpha_1 MDP_{it} + \alpha_2 DTP_{it} + \alpha_3 \ln(MFPR)_{i,t-1} + \alpha_4 \ln(CN)_{i,t-1} + w_{it}.$$

Subscript *i* refers to the state *i*, *t* refers to the month *t*, and \ln denotes natural logarithm. The error term, w_{it} , is specified to have zero mean and to be homoscedastic within a given state but

heteroscedastic across states. *MDP* is the amount of a state's production base on a fifteen-month basis entered into the diversion program in billions of pounds. It has the same value for each month between January 1984 through March 1985 and zero otherwise. Variable *DTP* is the annual base production deleted, by state, in billions of pounds under the *DTP*. The variable *DTP* is structured as a cumulative variable since the termination program was subdivided into three exit periods (4/86–8/86, 9/86–2/87, and 3/87–8/87). Producers could submit bids for any or all of the periods but had to exit during the period associated with the accepted bid. *DTP* is zero for all observations prior to 4/86. From 4/86–8/86 it is the state's base exiting during the first period. From 9/86–2/87 it is the sum of a state's first and second periods' base production exiting. From 3/87–8/87 it is the total base exiting during the termination program. From 9/87 forward *DTP* is zero.

A reasonable argument can be made that the *DTP* variable should not become zero after the exit period but continue to have nonzero values because exited farmers under the DTP could not resume production for five years after exiting. As discussed in the results section, this seemingly more orthodox specification gave inferior results. Price of slaughter cows, which has been included in prior studies (e.g., Kaiser, Streeter, and Liu; LaFrance and de Gorter) is not included in (1). In all specifications examined, this price had a positive and significant coefficient which is contrary to expectations. This likely results from the fact that both dairy cows and beef cows were in a liquidation phase of the cattle cycle from 1983–88, the sample period used for estimation and prediction.

PPC is a function of the *MDP*, lagged *MFPR*, monthly binary variables to account for the changes in range conditions and lagged *PPC*. It is written as

$$(2) \quad \ln(PPC)_{it} = \beta_0 + \sum_{j=1}^{11} \beta_j D_{jit} + \beta_{12} MDP_{it} + \beta_{13} \ln(MFPR)_{i,t-1} + \beta_{14} \ln(PPC)_{i,t-1} + v_{it},$$

where $D_{jit} = 1$ if the observation occurs in the *j*th month, 0 otherwise. Preliminary estimation indicated the *DTP* variable was insignificant at the .2 level for a one-sided test so it was not included in the final model. This is not surprising since whole herds were eliminated as opposed to selective culling. The error v_{it} term is assumed to have the same properties as w_{it} al-

though v_{it} and w_{it} are not restricted to have equal variances.

By specifying *MFPR* to enter in lagged form to account for expectations used by farmers to adjust current production, both equations can be estimated without additionally estimating a demand model. Kaiser, Streeter, and Liu use a similar justification in estimating these supply equations separately from the demand equations in their model.

Swamy, Conway, and LeBlanc give a number of reasons the conventional fixed coefficient regression model may not be an accurate specification for econometric estimation. In a model with cross-sectional data, it can be argued that all the regression coefficients being identical is much more unlikely than having them vary from one cross-sectional entity to another. For the particular case at hand, the alternatives to dairying across states are quite varied as are climate and geography, so it is reasonable to expect coefficients to differ. As a result, the responses to various government programs are likely to be different. These reasons lead us to select a Swamy RCR as a plausible stochastic specification of the model. A fixed coefficient model for each state estimated as seemingly unrelated regressions is an alternative specification. It is rejected because it would not yield a mean response vector, which could be used for assessing program impacts for states not included in the sample.

The basic Swamy RCR specification is given as

$$(3) \quad y_i = x_i\beta_i + e_i \quad e_i \sim (0, \sigma_{ii}^2 I),$$

where y_i contains T observations on the dependent variable, x_i is a $T \times k$ matrix of observations on k explanatory variables, and e_i is a vector of T observations on the error terms. The β_i are the sum of the mean coefficient vector, β and a random disturbance vector, μ_i . Thus,

$$\beta_i = \beta + \mu_i \quad \mu_i \sim (0, \Delta).$$

Further, it is assumed that $E(\mu_i e_j) = 0$ for all i, j and that μ_i and μ_j are uncorrelated for all $i \neq j$. The mean coefficient vector, β , is estimated by generalized least squares (GLS). Since the matrix Δ is unknown, it must be estimated. The procedure recommended by Judge et al. (p. 542) is used.

In estimating the impact of the two programs, the econometric model gives more accurate simulations by using predictions of the coefficient vectors for each state instead of the mean vector

for all states. State coefficient vectors are computed using the best linear unbiased predictor (*BLUP*) as given in Judge et al. (p. 541). The formula for the prediction of the i th coefficient vector is given as

$$(4) \quad \hat{\beta}_i = \hat{\beta} + \Delta x_i' (x_i \Delta x_i' + \sigma_{ii} I)^{-1} (y_i - x_i \hat{\beta}),$$

where $\hat{\beta}$ is the GLS estimator of β . However, since Δ is unknown, the *BLUP*'s are estimated by substituting the estimate for Δ .

Data and Estimates

Data for *CN* and *PPC* come from various issues of the USDA publication *Milk Production*. The data for the blend price of milk and the feed price of 16% dairy ration are from various issues of *Agricultural Prices* (USDA). The blend price is by state by month. The feed prices are by state, by month from January 1983 through February 1986. After 1986 they are reported on a regional basis for January, April, July, and October. Monthly observations for the nonreported months are obtained by interpolation. Data for the MDP come from Miller and Carman and data for the DTP come from the Agricultural Conservation and Stabilization Service (personal communication). The twenty-one states in the sample produced 85% of the commercial milk in the United States from 1983 through 1988. About 78% of the base production diverted in the MDP was accounted for by the twenty-one states as well as 75% of the DTP-enrolled base. Hence, the sample is assumed to give an accurate representation of mean producer response in the United States.

Estimates of the two models are displayed in tables 1 and 2.¹ The signs of all the coefficients are consistent with expectations. Somewhat surprising are the high levels of the coefficients of determination for the models, particularly the *CN* equation. In a cross section such as this, considerable variation arises because cow numbers vary substantially across states. With a lagged dependent variable in the equation and lack of substantial variation in cow numbers from month to month within a state, it is not surprising that the coefficient of determination is high.

¹ The blend price of milk is gross price. Producers paid the \$.15/cwt. assessment and various deductions due to the price support program and Gramm-Rudman-Hollings. The model was estimated with these values netted out of blend price, but there was little change in the estimates of the mean coefficients.

Table 1. Random Coefficient Regression Estimates of the Mean Coefficient Vector, *CN*

Regressor	Parameter Estimates	<i>t</i> -Ratios	$\delta_{ii}^{1/2a}$
Intercept	.1088	1.6757	.2660
<i>MDP_{it}</i>	-.0094	-2.4930** ^b	.0145
<i>DTP_{it}</i>	-.0206	-4.3821**	.0184
$\ln(MFPR)_{it-1}$.0101	2.4820*	.0152
$\ln(CN)_{it-1}$.9791	81.73**	.0494
$R^2 = 0.9992$	<i>PRMSE</i> = .3212	<i>MSE</i> = .0004	1,365 observations

Source: estimated.

^a The column under $\delta_{ii}^{1/2}$ is the estimated standard deviation of the *i*th coefficient, i.e., the square root of the *i*th diagonal element of the estimate of Δ .^b Single asterisk indicates significant at the 5% level of significance; double asterisk, significant at the 1% level of significance.

As mentioned earlier, an argument can be made that the *DTP* variable should maintain its value for those periods subsequent to the exit periods since those farmers exiting were precluded from engaging in dairying activity for five years after exiting. When *DTP* is so structured the percent root mean square error for the *CN* equation is 1.41, more than four times its value in table 1. In alternative specifications *MDP* in the post exit months was multiplied by d^r where d was varied from 0 to 1 and r was the number of months since the last exit month, 8/87. *PRMSE* was minimized, using a grid search, at $d = 0$. Thus, the conclusion must be that the mean behavior across states returned to the same behavioral relationships as before the *DTP*. Such behavior perhaps reflects those individuals not participating in the *DTP* expanding existing operations or new producers entering the dairy industry.

Two elasticities of supply can be computed from the model by multiplying the *CN* model by the *PPC* model to get total supply. The short-run supply elasticity is .0402, which is very inelastic. The long-run elasticity is .750, which is somewhat higher than those reported in other studies. Wilson and Thompson found an approximate long-term elasticity of .521, which is also close to the elasticities found by Halvorson for two of his five models. However, both of these studies used only time-series data which could lead to smaller estimates than ours. The elasticity estimates in the present study are sensitive to the fact that the coefficient of the lagged dependent variable is nearly one in the *CN* equation. With a monthly equation for cow numbers, it is not surprising that this parameter would be near to one since herd sizes do not vary much from month to month.

Table 2. Random Coefficient Regression Estimates of the Mean Coefficient Vector, *PPC*

Regressor	Parameter Estimates	<i>t</i> -Ratios	$\delta_{ii}^{1/2b}$
Intercept	.7981	5.430** ^b	.5848
<i>MDP_{it}</i>	-.0257	-2.388**	.0430
<i>D_{1t}</i>	-.0136	-2.394**	.0236
<i>D_{2t}</i>	-.0884	-11.38**	.0339
<i>D_{3t}</i>	.0673	10.14**	.0286
<i>D_{4t}</i>	-.0233	-2.165*	.0480
<i>D_{5t}</i>	.0202	2.034*	.0441
<i>D_{6t}</i>	-.0658	-5.668**	.0516
<i>D_{7t}</i>	-.0318	-2.690**	.0527
<i>D_{8t}</i>	-.0437	-3.980**	.0490
<i>D_{9t}</i>	-.0617	-7.312**	.0370
<i>D_{10t}</i>	-.0147	-2.652**	.0230
<i>D_{11t}</i>	-.0625	-9.602**	.0278
$\ln(MFPR)_{it-1}$.0301	4.219**	.0203
$\ln(PPC)_{it-1}$.8871	40.52**	.0878
$R^2 = 0.9492$	<i>PRMSE</i> = .4553	<i>MSE</i> = 0.0010	1,365 observations

Source: estimated.

^a The column under $\delta_{ii}^{1/2}$ is the estimated standard deviation of the *i*th coefficient, i.e., the square root of the *i*th diagonal element of the estimate of Δ .^b Single asterisk indicates significant at 5% level; double asterisk, significant at 1% level.

As a further test of the equations' accuracy, the monthly production in each state was simulated over the sample period. The observed values of the dependent variables were used for the lagged dependent variables. Total production was simulated with percent root mean square errors (*PRMSE*) ranging from 4.05 to 1.03 with eighteen of the twenty-one states below 2.0. The *PRMSE* for monthly production aggregated over the twenty-one states was .834. Because the policy simulations use predicted values of the endogenous variables over observations from 2/83–6/88, it is useful to consider the *PRMSE*'s for the model using predicted values of the lagged endogenous variables. Over these sixty-five observations the *PRMSE*'s ranged from 6.74 to 2.03. For the aggregate amount the *PRMSE* is 2.11.

The policy variables also have the expected signs and all are statistically significant at the .05 level. Because of the double logarithmic specification the coefficients can be viewed as elasticities of reductions per month. Thus, the MDP had a short-run elasticity of .0351 per month during the program and DTP has a short-run elasticity of $-.0206$. However, because the coefficients in tables 1 and 2 are means of the various coefficients across states, they are not necessarily reflective of the aggregate effect of changes in independent variables since the population levels in the various states differ. This effect is illustrated later in the policy simulations.

The last columns in tables 1 and 2 give the estimated standard deviations of the coefficients for their variation over states. Using the BLUP's as computed from (4), the coefficient of MDP in the *CN* equation ranged from $-.0303$ to $.0042$. In the *PPC* equation the coefficient of MDP ranged from $-.0904$ to $.0035$. The coefficient of DTP ranged from $-.0364$ to $-.0075$. Clearly, these rates have different short-term implications. However, it is more important to analyze the dynamic implications of the programs. In particular, the corresponding variations of the coefficients of the lagged dependent variables have a major influence on the overall impact of a policy in a particular state. The state-by-state variations in impact are considered in the simulation section.

As a final test of the model's usefulness, the last six months of 1988, which were not included in the estimation sample, were simulated. In this simulation the predicted values of the lagged endogenous variables were used instead of the observed values for every month but

July 1988, which used the observed values in June 1988. The *PRMSE*'s for total production ranged from 12.6 to 1.75 with a mean of 5.72. For the monthly aggregate production over twenty-one states for six months, the *PRMSE* was 5.69%. Given the various measures of goodness of fit, the model is considered useful for policy analysis.

Simulations and Policy Analysis

Because of the dynamic nature of the model, interpreting the coefficients of the binary variables as representing the impact of the policies can be misleading. To avoid this, the model was simulated four different ways from February 1983 through June 1988. In the base simulation it was assumed that both of the programs were in effect. Then the model was simulated assuming that only the MDP was in effect, then only the DTP, and then with neither program in effect. The impact of the programs can be estimated by comparing the production levels in the simulations with the policy variables at nonzero levels with production levels assuming no programs. Comparisons are made on an aggregate level as well as state by state. This latter comparison is important to measure the heterogeneity of state responses to per hundredweight participation in the programs.

In the policy simulations deleting one or both of the programs, blend prices are assumed to be endogenous with feed prices at their observed levels. The endogenous blend prices are computed in the following way. Recall that milk output levels in period t are determined as a function of prices lagged one period. The production levels in period t with and without the program(s) are computed. It is assumed that all additional production due to absence of a program(s) would be used in manufacturing grade.² Thus, the blend price is changed by computing the new proportions of fluid milk and nonfluid milk. This new blend price becomes the price used in the next period.

Subtracting production in the base simulation from the third simulation gives the impact of both of the programs taken together. These figures are shown under the column "COMBINE" in table 3. For the twenty-one states, the

² Manufacturing grade is used instead of Class II because a consistent Class II price by state is not available. In states where manufacturing prices are not available, Class III prices for milk market order districts were used.

Table 3. Impact of MDP and DTP on Milk Supply, by State, 1983, 2–1988, 6

State	Milk Reduction					
	MDP ^a	BMDP ^b	DTP ^c	COMBINE ^d	MARGIN ^e	BDTP ^f
	(billion pounds)					
CA	.649	.892	1.699	2.534	1.685	1.648
FL	.229	.396	.139	.567	.138	0.258
ID	.439	.205	.232	.670	.231	0.461
IL	.076	.224	.186	.264	.188	0.155
IN	.453	.161	.278	.720	.267	0.159
IA	.823	.400	.646	1.455	.632	0.309
KY	.740	.313	.267	.995	.255	0.182
MD	.065	.053	.125	.188	.123	0.117
MI	1.174	.282	.601	1.753	.579	0.586
MN	.660	.781	1.900	2.552	1.892	0.870
MO	.826	.390	.315	1.134	.308	0.371
NC	.171	.084	.220	.386	.215	0.204
NY	.733	.367	1.175	1.915	1.182	0.437
OH	.748	.267	.450	1.134	.436	0.242
PA	.757	.254	.341	1.095	.338	0.253
TN	.131	.246	.109	.240	.109	0.166
TX	1.115	.475	.533	1.625	.510	0.602
VA	.179	.126	.158	.335	.156	0.186
VT	.204	.112	.081	.235	.081	0.156
WA	.366	.161	.603	.954	.598	0.501
WI	-.214	1.07	2.768	2.548	2.762	0.714
Total	10.32	7.259	12.826	23.009	12.685	8.577

^a MDP denotes milk reduction due to MDP only (assumes DTP was not implemented).

^b Amount of base production enrolled for diversion into MDP, 15-month basis (Miller and Carman)

^c DTP denotes milk reduction due to DTP only (assume MDP was not implemented)

^d COMBINE denotes milk reduction due to both MDP and DTP.

^e Margin denotes the marginal impact of the DTP in addition to MDP. Computed as COMBINE less MDP.

^f Amount of annual base production by herds liquidated in DTP (ASCS, personal communication).

two programs resulted in a cumulative reduction of milk production of 23 billion pounds through June of 1988. Of this the MDP generated 10.32 billion pounds.

It would be somewhat misleading to say that the DTP resulted in a cumulative 12.83 billion pound decrease through June 1988 because the column under DTP assumes that the DTP was implemented without the MDP. The marginal contribution of the DTP is 12.68 billion because of when it followed the MDP. The fifth column in table 3 is computed as the values of the fourth column less the corresponding values of the first column. There is a 1.1% decrease in the reductions resulting from the DTP because of when it was implemented relative to the MDP. Presumably, if the DTP had begun sooner after the conclusion of the MDP, the DTP's impact would have been even less.

Long- and short-term aspects of the two programs are clear from figure 1. Twelve months after the cessation of the MDP, production levels had recovered almost half of the reduction due to the MDP. However, the impact of the DTP was greater than the MDP on a percentage of total output basis and the rate of recovery from

the DTP was slower than from the MDP. Hence the DTP appears to have longer-term effects than the MDP.

Effectiveness of the programs varied by state. Participation rates varied for both programs by state and so did diminution of production per pound of base production enrolled. Columns 2 and 6 in table 3 show the amount of production enrolled. Dividing column 1 by 2 and column 5 by 6 gives the ratios of production declines to base enrolled. For the MDP these ranged from 4.16 for Michigan to -.2 for Wisconsin. The ratio for all twenty-one states combined was 1.42. For the DTP the ratios ranged from 3.88 (Wisconsin) to .503 (Idaho). The ratio for the twenty-one states combined was 1.48. However, these figures account for reductions only through June 1988. The DTP ratios will increase as the impact of the DTP is felt beyond June 1988.

Concluding Comment

Swamy random coefficient regression equations with cow numbers and production per cow as dependent variables were estimated using a time

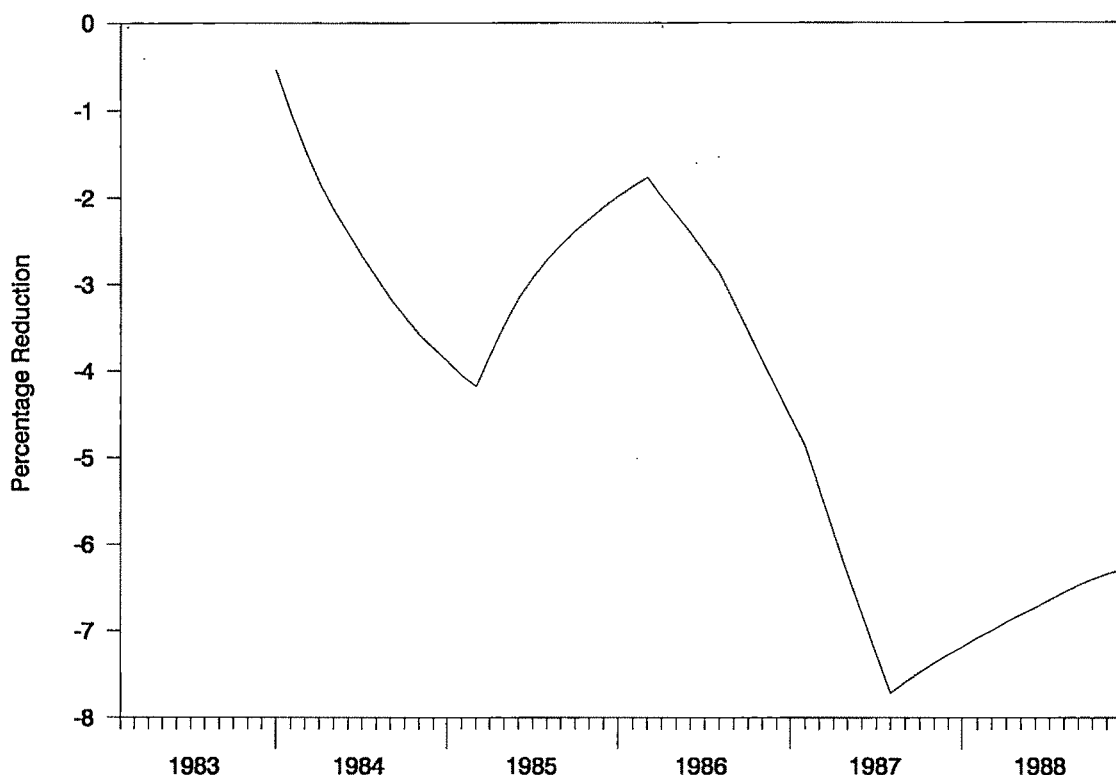


Figure 1. Percentage reduction in milk production due to DTP and MDP, 21 states

series of cross-sectional data on twenty-one major milk-producing states in the United States. These equations were used to estimate the impact of two very different programs that were implemented in the mid-1980s to reduce surplus commercial milk production. The first program was the milk diversion program which had a short-term impact. Within a year after its cessation, production had nearly recovered half of the decrease induced by the MDP. On the other hand, the whole herd buyout ceased in August 1987, but the rate of recovery to previous levels was slower. Hence, the dairy termination program has had a longer-term impact than the milk diversion program. However, the five-year ban on production by producers exiting under the DTP appears to have had little, if any, effect. Herd size began increasing immediately after the last exit month to pre-DTP equilibrium levels.

The effect of the programs was not uniform across the states. This is indicated by variations in coefficients across states as well as by the impacts per unit of base enrolled in the two programs. An explanation for this variation is that in some states the demand for milk was probably sufficiently strong that nonparticipating

producers increased production substantially. For future dairy supply management programs, an important implication is that the same program could have very different regional impacts, perhaps requiring coordination to obtain desired effects.

[Received October 1989; final revision received October 1990.]

References

- Berry, C. R., and L. V. Blakley. "Adjustments in Milk Production Under the Milk Diversion Program." University of Arkansas Agr. Exp. Sta. Bull. 899, 1987.
- Chen, Dean, Richard Courtney, and Andrew Schmitz. "A Polynomial Lag Formulation of Milk Production Response." *Amer. J. Agr. Econ.* 54(1972):77-83.
- Halvorsen, Harlow W. "The Response of Milk Production to Price." *J. Farm Econ.* 40(1958):1101-13.
- Judge, George G., William E. Griffith, R. Carter Hill, and Tsoung-Chao Lee. *The Theory and Practice of Econometrics*. New York: John Wiley & Sons, 1980.
- Kaiser, Harry M., Deborah H. Streeter, and Donald D. Liu. "Welfare Comparisons of U.S. Dairy Policies with and without Mandatory Supply Control." *Amer. J. Agr. Econ.* 70(1988):848-58.
- Kaiser, Harry M., and Loren W. Tauer. "Impact of Bovine

- Somatotropin on U.S. Dairy Markets Under Alternative Policy Options." *N. Cent. J. Agr. Econ.* 11 (1989):59-74.
- LaFrance, J. T., and H. deGorter. "Regulation in a Dynamic Market: The U.S. Dairy Industry." *Amer. J. Agr. Econ.* U.S. Dairy Industry." *Amer. J. Agr. Econ.* 67(1985):821-32.
- Matthews, Dick. "Dairy Notes." Snohomish County (Washington) Coop. Extens. Serv., April 1986.
- Miller, James J., and Clifford M. Carman. "Participation in the Milk Diversion Program." *Dairy Situation*. Washington DC: U.S. Department of Agriculture, Econ. Res. Serv. DS376, March 1984.
- Swamy, P. A. V. B. *Statistical Inference in Random Coefficient Regression Models*. Berlin, New York: Springer-Verlag, 1971.
- Swamy, P. A. V. B., Roger K. Conway, and Michael R. LeBlanc. "The Stochastic Coefficients Approach to Econometric Modeling. Part I: A Critique of Fixed Coefficients Models." *J. Agr. Econ. Res.* 40(1988): 2-10.
- Whipple, Glen D., Charles Powe, and Morgan-Gray. "An Economic Analysis of Selected U.S. Dairy Program Changes." *S. J. Agr. Econ.* 17(1985):181-91.
- Wilson, Robert R., and Russell G. Thompson. "Demand, Supply, and Price Relationships for the Dairy Sector, Post-World War II Period." *J. Farm Econ.* 49 (1967):360-71.

Determinants of Agricultural Economics Faculty Retirement

Josef M. Broder, Fred C. White, and Teresa D. Taylor

The retirement experiences of faculty who had retired from 1862 land grant universities are described. Health, financial, family, and professional characteristics of early retirees are contrasted to those of late retirees. Determinants of faculty retirement are identified in the context of a general retirement model. Model results indicate that years-of-service to retirement programs, number of children, mental health, health trends, age at initial employment, and faculty-spouse age differences are positively related to retirement age. Salary levels, net worth, and extension appointments are inversely related to retirement age. Implications of recent changes in the agricultural economics profession on expected retirement behavior are discussed.

Key words: faculty compensation, faculty retirement, professional affairs, retirement profiles.

The typical agricultural economist pursues a productive career for thirty to forty years and then leaves the profession through retirement. Given life expectancies that exceed eligible retirement age, most of our profession will join the ranks of the retired. Among the major life-cycle decisions made by agricultural economists, few are as difficult as the decision to retire. Despite the inevitability of retirement, little is known about retired agricultural economists. While the retirement literature offers some insights into the retirement experience of university faculty, the mission and appointment structure of agricultural economics faculty are unique and worthy of separate analysis.

In response to this absence of information on retired agricultural economists, a survey of agricultural economics faculty who had retired from 1862 land grant universities was conducted. The purpose of the survey was to develop personal profiles of retired agricultural economists and to gain a better understanding of the decision to retire. The objectives of this paper are (a) to describe biographical characteristics of retired faculty, (b) to describe the health and financial pro-

files of retired faculty, and (c) to identify factors that influenced the decision to retire.

The agricultural economics profession has maintained an ongoing interest in professional affairs. Recent studies have examined rankings of individual departments (Beilock, Polopolous, and Correal), assessments of journals used by agricultural economists (Broder and Ziemer 1985), the journal publication process (Brorsen), productivity of agricultural economics faculty (Broder and Ziemer 1982), undergraduate student affairs (Litzenberg and Schneider, Erven), extension-related issues (McDowell; Knight, Johnson, and Finley) and professional interests (Pope and Hallam). These publications offer useful insights into the life cycle of our profession. Largely excluded from this literature are studies on the retirement experiences of faculty. In their absence, the life-cycle assessment of the profession would be incomplete.

Current Study

Data for the current study were taken from a survey of retired agricultural economics faculty at 1862 land grant universities, conducted in the spring of 1988. Department heads at forty-eight 1862 land grant universities were asked to provide names and addresses of living faculty who had retired from their departments. The 404 names received from 42 of the department heads responding to the request were used as the sur-

Josef M. Broder, Fred C. White, and Teresa D. Taylor are professors, Department of Agricultural Economics, University of Georgia, and an agricultural economist, Tennessee Valley Authority.

The authors would like to thank Joseph P. Walsh, Jr., of the Social Security Administration, and Marc A. Eason, Agricultural Business Office, University of Georgia, for their many useful comments on this paper.

vey population and were mailed questionnaires. Two hundred nine (209) usable responses were received for a response rate of 51.7%.

Respondent Characteristics

Respondents were asked to describe their professional activities and biographical status at the time of retirement (table 1). These characteristics were used to develop profiles of early and late retirees. First, working definitions of early and late retirement were developed for this analysis. Second, mean differences between a particular retiree group and all other groups were tested using Student's *t*-statistic.

While definitions of early and late retirement are related to particular retirement systems, the literature draws a clear distinction between early and late retirees with the normal age of retirement at 65 (Bazzali). Because the average retirement age of agricultural economics faculty was 63, with 71% retiring before age 66, this

delineation was deemed too restrictive for this study. To account for the earlier retirement age of agricultural economics faculty and the flexibility of many faculty retirement programs, the period of normal retirement was expanded to include ages 62 through 65. As in previous studies, respondents retiring after age 65 were described as having retired late. Respondents retiring before age 62 were described as having retired early. This definition of early retirement is consistent with years of eligibility for retirement benefits as defined by the Social Security Administration. Persons fully insured under Social Security can retire as early as age 62 with permanent reductions in retirement benefits or at age 65 with full benefits (U.S. Department of Health and Human Services). Approximately 77% of respondents had made contributions to Social Security and were eligible for Social Security benefits.

Respondents were asked to indicate their faculty appointments at retirement. About half of the respondents had multiple appointments. Ex-

Table 1. Average Characteristics of Retired Agricultural Economics Faculty, 1988

Characteristic at Retirement	Age at Retirement		
	<62	62-65	>65
Group delineation	Early	Normal	Late
Number of observations	62	76	60
Faculty rank:	(%)		
Professor	72.6	71.1	80.0
Associate professor	16.4	15.8	10.0
Assistant professor	2.7	2.6	3.3
Other	8.2	10.5	6.7
Field of study:			
Production/finance	39.1	40.5	31.0
Marketing/policy	27.5	31.1	46.6
Resources/development	8.7	13.5	15.5
All other	24.6	14.9	6.9
Faculty appointment: ^a			
Teaching	38.4	48.7	60.0
Research	34.3	50.0	53.3
Extension	71.2	42.1	40.0
Administration	15.1	25.0	26.7
Other	2.7	4.0	3.3
Multiple appointments	56.2	48.7	49.3
Female	1.3	1.3	1.7
Family status:			
Married	90.4	94.7	85.0
Single	5.5	2.6	8.3
Widowed	2.7	2.2	5.0
Divorced	1.4	0.0	1.7
Reason for retiring:			
Mandatory	6.8	19.1	54.0
Financial	37.3	22.2	12.0
Health	30.5	41.3	22.0
Other	25.4	17.5	12.0

^a Percent of respondents with appointments.

tension appointments were most frequent among respondents, followed by teaching and research. The early retirement group was dominated by extension faculty with over two-thirds of this group having extension appointments. The late retirement group was dominated by faculty with teaching appointments, followed by faculty with research appointments. These data suggest that teaching and research faculty tended to retire later than their extension counterparts. Financial considerations were cited as the major reason for retirement among early retirees, while health reasons dominated those in the normal age of retirement. As expected, mandatory retirement was most common among late retirees.

The proportion of respondents with extension appointments was found to be higher than in the working population of agricultural economics faculty (Broder and Ziemer 1982, 1985). These differences were attributed to the research design and appointment-related differences in retirement age. Because extension faculty tended to retire early, the percentage of retired faculty with former extension appointments should exceed the percentage of extension faculty among the working population of agricultural economics faculty. Although no systematic tests were

made to identify response bias, the sample population was thought to represent the population of retired agricultural economics faculty.

Retirement Profiles

Retirement profiles of respondents are shown in table 2. The average years of birth for the early, normal, and late respondent groups were 1923, 1917, and 1912, respectively. Age differences between respondent and spouse appeared to increase with retirement age from 2.5 years for early retirees to 4.0 for late retirees. No significant differences were found in the percentage of spouses employed at the time of the respondent's retirement. However, early retirees tended to retire before their spouses' retirement, while late retirees tended to work beyond their spouses' retirement. The fewest children were reported by late retirees, while early retirees reported having the largest number of persons living in the current household.

Respondents who retired early reported having the fewest years of service in university retirement programs and the most years in federal retirement. Late retirees reported just the op-

Table 2. Average Retirement Profiles of Retired Agricultural Economics Faculty, 1988

Characteristic at Retirement	Age at Retirement		
	<62	62-65	>65
Year of birth			
Respondent	1922.7***	1917.3	1912.2***
Spouse	1925.2***	1920.8	1916.3***
Year of retirement			
Respondent	1980.6	1981.2	1979.9
Spouse	1981.9	1981.4	1978.2
Spouses employed at respondents' retirement (%)	46.6	40.8	43.3
Number of children	3.3	2.8	2.5*
Number of persons in current household	2.4	2.0	2.0
Number of job changes	2.7	2.3	3.0
Respondent's years of service toward ^a			
University retirement	16.4***	22.7	24.3**
Social Security	11.6***	19.6*	20.3**
Federal retirement	22.6***	12.1*	8.4***
Other programs	7.0*	5.0	3.3*
Spouse's years of service toward			
University retirement	2.1	1.5	0.8
Social Security	7.1	7.7	7.2
Federal retirement	0.7	0.6	0.3
Other programs	3.7	1.0***	3.8

^a Means different at alpha levels = 0.01 (***), 0.05 (**), and 0.10 (*).

^b The percentage of respondents contributing to individual retirement programs were: university 76.6%, Social Security 76.6%, federal 53.1%, and other 32.0%.

posite. These data suggest that for the survey population, retirement age appears to be inversely related to years of service in federal retirement programs and directly related to years of service in university retirement programs. Spouses had more years of service in Social Security than in all other retirement programs combined.

Health Profiles

Previous studies have shown health to be a primary determinant in the decision to retire (Anderson and Burkhauser). The sources of health data used in the literature include self-reported health and health based on mortality experience. The current study attempted to measure self-reported health in a cross-sectional framework. Self-reported health data were obtained by asking respondents to rate their physical health and mental well-being before and after retirement. Respondent ratings were based on a seven-point scale with a value of 7 = excellent, a value of 4 = average, and a value of 1 = poor.

Self-reported health data are shown in table 3, including (a) the average health for the 1970–88 period, (b) the change in health during the

1970–88 period, and (c) health at retirement. The lowest average levels of physical and mental health were reported by respondents who retired early. Retirees from all categories reported that their physical health had declined during the study period. The mental health of all but the late retirement group had improved during the study period. Significantly higher levels of mental health at retirement were reported by late retirees. The lowest levels of physical and mental health at retirement were reported by early retirees.

To learn more about changes in self-reported health, respondents were asked to list special reasons for changes in their physical and mental health. Early retirees reported stress as the most frequent physical ailment, followed by cardiovascular, and a variety of other disorders. Early retirees cited mental stress and declining physical health as major reasons for changes in mental well-being.

Two-thirds of late retirees gave no specific reasons for changes in their physical health. Among reasons given by this group were cardiovascular and other physical health problems. Declining physical health and mental stress were cited by late retirees as major reasons for changes in mental health. When contrasted to early re-

Table 3. Average Physical and Mental Well-Being of Retired Agricultural Economics Faculty, 1988

Characteristic	Age at Retirement		
	<62	62–65	>65
Physical health	(index ^a)		
Average 1970–88	5.29 ^{ab}	5.71	5.65
Change 1970–88	–0.53	–0.47	–0.57
At retirement	5.33 [*]	5.67	5.73
Mental health			
Average 1970–88	5.60 ^{**}	6.08	6.06
Change 1970–88	0.15	0.13	–0.23 [*]
At retirement	5.45 ^{***}	5.95	6.09 [*]
Reasons for change in physical health	(%)		
None	42.25	51.35	66.07
Cardiovascular	15.49	18.92	12.50
Stress	16.91	5.41	5.36
Cancer related	7.04	4.05	5.36
Other	18.31	20.27	10.71
Reasons for changes in mental health			
None	43.48	60.53	63.16
Declining physical health	14.49	10.53	14.04
Mental stress	28.99	13.16	12.28
Change in lifestyles	5.80	10.53	8.77
Family problems	7.25	5.26	1.75

^a Self-reported data based on an index of 1 to 7, where 7 = excellent and 1 = poor.

^b Mean different at alpha levels = 0.01 (***), 0.05 (**), 0.10 (*).

tierees, late retirees reported having fewer physical and mental health problems.

Since health was thought to influence the decision to retire, health indexes were also presented by respondent age. Self-reported health, shown in figure 1, suggests that mental and physical health of respondents were similar through age 55 and diverged afterwards. After age 55, respondents reported mental health as being superior to physical health with the greatest divergence occurring at ages 70 to 75. Conversely, the variation in reported physical health was greater than that for mental health with the greatest divergence also occurring at ages 70 to

75. Discrepancies between mental and physical health suggest that physical health alone may not be a reliable indicator of respondent well-being or predictor of faculty retirement.

Financial Profiles

Previous studies have shown that the financial status of workers has a significant influence on their decision to retire (Fields and Mitchell). Hence, respondents were asked to report retirement incomes, expenses, and net worth. Financial profiles of retirees are shown in table 4.

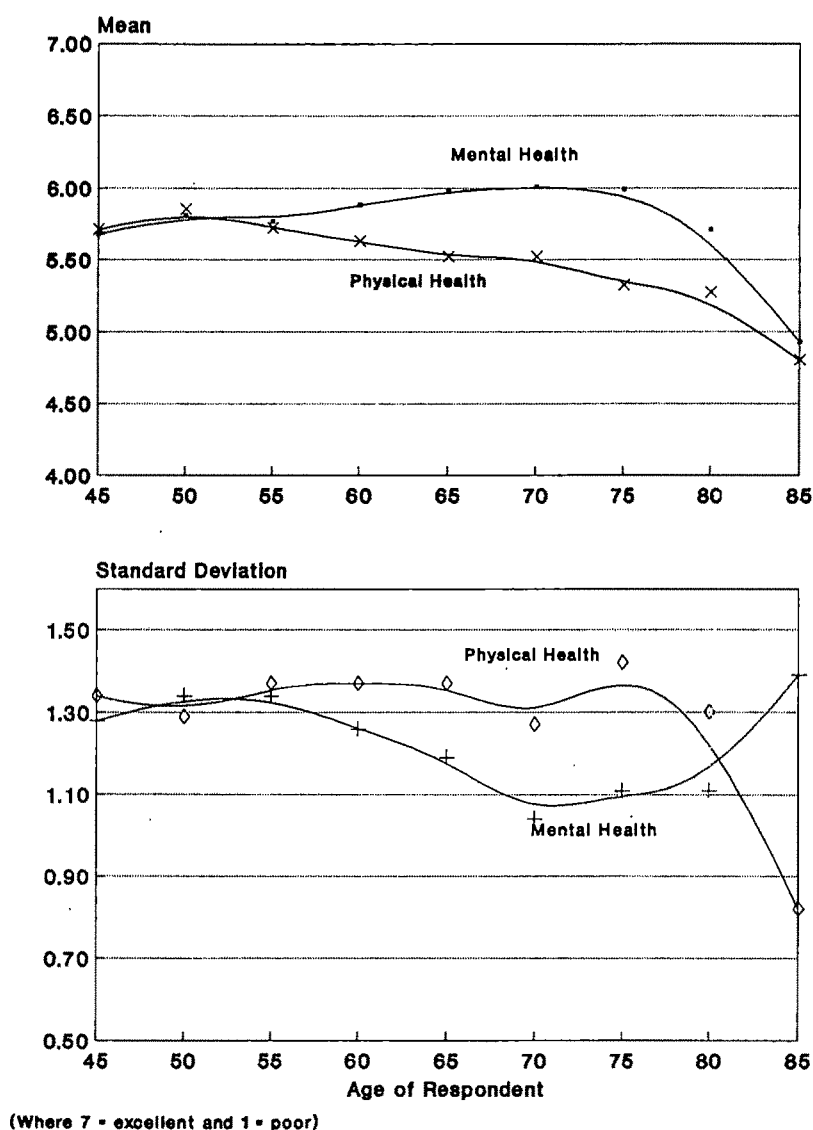


Figure 1. Health profiles of retired agricultural economics faculty

Table 4. Average Financial Status of Retired Agricultural Economics Faculty, 1988

Characteristic	Age at Retirement		
	<62	62-65	>65
	----- (\$1988) -----		
Respondent's annual salary during last year of employment	55,066	52,094	53,417
Spouse's annual salary during respondents last year of employment	10,454***	4,611*	5,082
Respondent's net worth			
Home	100,301	91,500	105,172
Other investments	241,068	253,170	202,107
Current annual income from all pensions			
Respondent's	33,231	34,821	33,648
Spouse's	3,256	3,681	4,418
Current annual income from all other sources			
Respondent's	18,232	17,196	20,686
Spouse's	6,144	5,222	5,926
Growth in retirement income minus cost-of-living (%)	1.7	1.0	2.6
Purchased life insurance to maximize value of pension (%)	17.8	18.4	10.0
Hours per week on professional activities			
Compensated ^b	7.7**	4.6	3.3
Uncompensated	3.4**	5.0	7.7

^a Means different at alpha levels = 0.01 (***), 0.05 (**), and 0.10 (*).

^b Percentages of respondent groups with 20 or more hours of compensated professional activities were early 12.6%, normal 4.0%, and late 5.0%.

Significant differences were found among salaries earned by spouses during the respondent's last year of employment. Spouses of early retirees earned significantly more than spouses of later retirees.

Respondents were asked to report their net worths in two categories: home and other investments. The largest net worth in home was reported by late retirees, while the largest net worth in other investments was reported by early retirees. However, differences in net worths were not statistically significant. No significant differences were found in the annual incomes received from all pensions held by the respondent and spouse.

Early retirees reported earning \$55,066 in their last year of employment, adjusted to 1988 dollars by the consumer price index. The spouse's salary at the time of the respondent's retirement was reported to be \$10,454 in 1988 dollars. Thus, combined earnings of early retirees and their spouses were \$65,520. When these combined preretirement earnings were compared to total postretirement income (pensions and other

sources) of \$60,863, the gross earnings of early retirees decreased by 7.1% after retirement. By contrast, gross earnings of normal and late retirees increased 7.4%, and 10.6%, respectively, after retirement. These data suggest that monetary costs of retirement appear to decline with retirement age.

These financial data also suggest that postretirement incomes from nonpension sources are important. These findings are consistent with respondents' assessment of how well their retirement incomes have kept pace with their cost of living. Despite observed postretirement reductions in gross earnings, early retirees indicated that the growth in their retirement income had exceeded cost-of-living increases by 1.7%. Comparable growth patterns were reported by late retirees, while retirees in the normal group reported smaller gains in real income after retirement. This growth in postretirement income was attributed to cost-of-living adjustments in Social Security payments as shown in table 2 and asset disinvestment strategies of retirees as shown by declining net worths in figure 2. Some

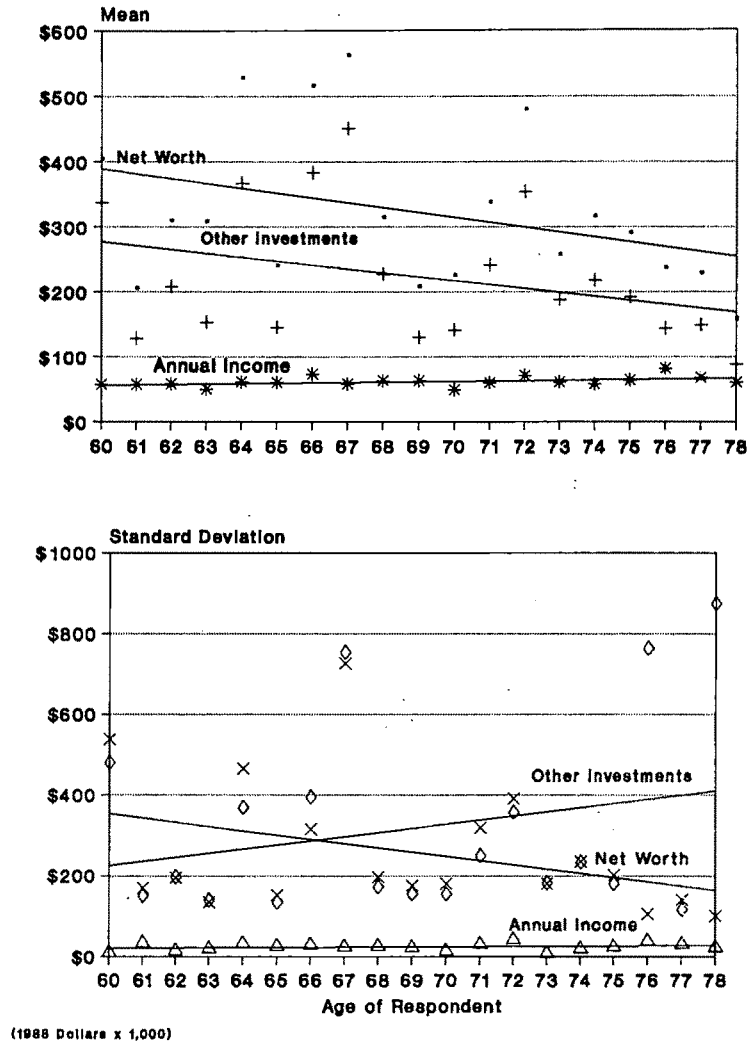


Figure 2. Financial profiles of retired agricultural economics faculty

of this growth in retirement income was also attributed to income from professional activities. Early, normal, and late retirees reported spending 7.7, 4.6, and 3.3 hours per week, respectively, on compensated professional activities.

A comparison of financial profiles by age of respondent is shown in figure 2. Linear trends of mean retiree incomes and net worth reveal constant incomes and gradual declines in net worths. Given that home values and annual income streams remain constant, retirees appear to be disinvesting other investments to maintain their annual incomes. These data suggest that financial profiles of retired faculty are likely to be affected by the length of the anticipated retirement period. Lengthy retirement periods would necessitate greater net worths and disin-

vestment over a longer time frame. Respondents anticipating longer retirement periods were more likely to purchase life insurance to maximize the value of their pensions and provide financial security for their spouses. Lengthy retirement periods were also associated with declines in the variability of total net worth and increases in the variability in other investments.

Retirement Model

Relationships among retiree health, finances, and biographical characteristics, and retirement age were examined in a general retirement model. The purpose of this retirement model was to identify and estimate determinants of retire-

ment. The conceptual framework upon which the model was developed assumes that faculty maximize utility subject to various monetary and nonmonetary constraints. Faculty derive utility from combinations of work and leisure, subject to the constraints imposed by income, health, and family obligations. The relative prices of work and leisure determine the number of hours the respondent works each period. Likewise, the relative prices or opportunity costs of pre- and post-retirement influence faculty retirement decisions, subject to institutional and market constraints.

Work and leisure are assumed to be imperfect substitutes in consumption, suggesting that retirement is a rational alternative to the faculty's career. Imperfect substitution between these activities suggest that the transition to retirement is apt to be gradual, with the faculty member working fewer hours near retirement and maintaining some professional activities after retirement (Honig and Hanoch).

The decision to retire is based on the faculty member's expected capacity to enjoy retirement. Financial assets are assumed to enhance the utility derived from retirement, while poor health diminishes the ability to work or to derive utility from retirement. Family relationships can either enhance or diminish the relative utility from retirement. Family members can offer financial and personal support to retirees and make their retirement experience more enjoyable. Finally, the structure of retirement programs affects the relative costs and, thus, the age of retirement.

The faculty member's age at retirement was used as the dependent variable in an econometric model to explain retirement behavior among agricultural economics faculty. Explanatory variables suggested by the conceptual framework and previous research are as follows.

End-of-career financial profiles have been found to influence the decision to retire (Anderson and Burkhauser, Fields and Mitchell, Honig and Hanoch, Bazzali). For this study, five variables were used to describe financial profiles of respondents. Respondent's salary during last year of employment (*RSAL*) was entered in 1988 dollars as a measure of the opportunity cost of retirement and the potential earnings from university pensions. The inclusion of *RSAL* assumed that faculty retirement follows the attainment of a threshold level of postretirement income. Higher salaries are assumed to provide greater financial security during retirement. Because pensions are generally based on earnings before retirement, an inverse relationship was

expected between retirement age and respondent earnings, indicating higher salaries contribute to earlier retirement.

Because almost 90% of respondents were married, spousal earnings were also expected to influence the decision to retire. The annual salary earned by the respondent's spouse during the respondent's last year of employment (*SSAL*) was entered into the model in 1988 dollars and was expected to have a negative influence on retirement age. With few exceptions, retired faculty members had contributed to several retirement programs. Hence, the decision to retire was hypothesized to be based on contributions made by respondent and spouse to all retirement programs. Years of service to all retirement programs (*TOTSER*) were entered into the model on the assumption that contributions to retirement programs enhance the financial security of the retired household. The variable *years of service* was expected to have a positive impact on retirement age.

Other sources of retirement income were also hypothesized as important determinants of retirement. Of the financial assets owned by faculty, net worth in non-home investments was considered the more liquid and amenable to disinvestment. Disinvestments of non-home assets provide income to supplement employer pensions and increase financial security during retirement. Thus, non-home assets were entered into the model (*INVEST*) and expected to have a negative impact on retirement age.

Children serve as a potential source of financial and psychological support during retirement. Retirement also affords the satisfaction of spending leisure time with children and grandchildren. Thus, the number of children was entered into the model and expected to have a negative impact on retirement age.

Previous research has shown health is a primary determinant of retirement (Anderson and Burkhauser, Bazzali). This study hypothesized that respondent attitude and assessments about health just before retirement were instrumental in the retirement decision. Thus, self-reported mental health before retirement (*MENTAL*) was entered into the model and expected to have a positive impact on retirement age. Faculty who gave high ratings to their mental well-being during employment were expected to retire later.

Previous research has shown that morbidity rates have an influence on the retirement decision (Anderson and Burkhauser). Lacking life expectancy data on individual respondents, the year of retirement (*TREND*) was entered into the

model as a proxy for increases in general health care and life expectancies that have occurred during the study period. The availability of improved health care and greater life expectancies were hypothesized to have a positive impact on retirement age.

The final set of explanatory variables were included to account for and, thus, control for situational differences in retirement programs and age differences between respondent and spouse (Dorfman). The diversity of retirement programs across universities precluded systematic classification and testing of program differences. However, the observed differences in retirements by faculty appointments suggested that faculty with extension appointments tended to retire earlier than their counterparts in teaching and research. The percentage of the respondent's appointment designated as extension (*SER*) was entered into the model and expected to be inversely related to retirement age. Reasons for early retirement among extension faculty were attributed to the more generous retirement benefits and options available to extension faculty under the pre-1983 Civil Service Retirement System (CSRS). Under this system, extension faculty with 30, 20, and 5 years of service, respectively, could retire at ages 55, 60, and 62 without penalty (U.S. Office of Personnel Management). Faculty under CSRS were also exempt from Social Security taxes and were not subject to the more restrictive eligibility criteria for Social Security benefits (U.S. Department of Health and Human Services). With an average of 25.6 years of service to the Civil Service Retirement, many extension faculty were eligible to retire before age 62. The tendency for research and teaching faculty to retire later was attributed to more restrictive retirement benefits among university retirement programs (Eason).

Other factors associated with early retirement among extension faculty include greater job-related stress and greater opportunities for consulting. In cross-tabulations of faculty by majority appointment (>50%), the percentage of extension faculty reporting physical and mental stress was 13% and 28%, respectively. These levels were twice that reported by other faculty. Extension faculty also reported an average of 6.7 hours of compensated professional activities after retirement as compared with 4.4 hours reported by other faculty.

The value of a university retirement program is considerably enhanced once an employee becomes vested. Once vested, the value of the pension is increased substantially by matching

funds provided by the employer. Vesting often requires that an employee make contributions to a retirement program for a minimum number of years. The value of a pension, also depends on the total years of service. For example, thirty years of service are required by some programs to be eligible for full pension benefits. Both of these retirement program attributes are sensitive to the age of initial employment with the university. The younger one begins employment, the sooner these program requirements can be met and the sooner full benefits can be received at retirement. Thus, the respondent's age at initial employment with the university (*AGEURP*) was entered in the model and expected to have a positive impact on retirement age.

Retirement programs also differ in their policies on mandatory retirement. To capture these differences across respondents, a binary variable for mandatory retirement (*MAND*) was entered (Leigh). *MAND* was equal to one if the respondent retirement was mandatory, zero otherwise.

Finally, age differences between respondent and spouse were also thought to affect the retirement decision. Because retirement affords the individual or household an opportunity to travel and relocate, the respondent's decision to retire was hypothesized to be closely related to the spouse's decision to retire. Age differences between respondent and spouse (*AGEDIFF*) were entered into the model with the expectation that respondents with younger spouses tended to postpone retirement to accommodate their spouses' working career.

Results

Ordinary least squares estimates of the retirement model are shown in table 5. With one exception, all estimated parameters were significant and consistent with a priori expectations. The empirical model found that total service to all retirement programs, self-reported mental health, health care trends, age at initial employment with the university, mandatory retirement, and age differences between respondent and spouse were positively related to retirement age. End-of-career salaries, net worth in non-home investments, number of children, and extension appointments were inversely related to retirement age. The lack of significance with the net worth variable was attributed to suspected inconsistencies among respondents in reporting the value of pensions as non-home assets. Model

Table 5. Factors Associated with Retirement Age of Agricultural Economics Faculty, 1988

Variable	Variable Description	Mean	Coefficient
Dependent Variable <i>RETAGE</i>	Retirement age	63.00	-146.90 (111.68) ^a
Explanatory Variables <i>RSAL</i>	Respondent's annual salary at retirement (1,000's)	53.50	-0.05*** (0.03)
<i>SSAL</i>	Spouse's annual salary at retirement (1,000's)	6.77	-0.05** (0.02)
<i>TOTSER</i>	Total years service to all retirement funds	57.78	0.03* (0.02)
<i>INVEST</i>	Net worth in non-home investments (100,000's)	2.35	-0.04 (0.08)
<i>CHILD</i>	No. of children	2.89	-0.29* (0.17)
<i>MENTAL</i>	Index of self-reported mental health	5.82	0.33* (0.19)
<i>TREND</i>	Retirement year	1980.63	0.10* (0.06)
<i>SER</i>	Percent service appointment	40.10	-0.03*** (0.01)
<i>AGEURP</i>	Age at initial employment with university	41.93	0.06*** (0.02)
<i>MAND</i>	Binary = 1 if mandatory retirement; 0 if otherwise	0.26	3.70*** (0.61)
<i>AGEDIFF</i>	Respondent's age minus spouses's age	3.16	0.19*** (0.06)

$R^2 = 0.42$; number of observations = 182

^a Standard error shown in parentheses.

^b Significant at alpha levels = 0.01 (***), 0.05 (**), or 0.10 (*).

estimates lend support to the utility maximization framework of the retirement decision.

The implications of model estimates must be interpreted in relation to the magnitude of their impacts on retirement age. For example, faculty with 100% extension appointments tended to retire three years earlier than teaching and research faculty with no extension appointments. Each child hastened the retirement decision by four months. In successive years of the survey period, the average age at retirement increased by five weeks. An additional \$10,000 in household income hastened the retirement decision by six months. Faculty elected to retire 3.7 years before mandatory retirement.

Conclusions

The findings of this paper suggest that the decision to retire is complex and not based on any single criterion. Although many determinants of

retirement were identified, the impacts of many of these determinants are small in isolation. These findings of this study have implications for the inevitability of retirement and relatively narrow time frame during which the retirement decision is made. The inevitability of retirement should alert faculty and universities to take active roles in planning for retirement, both mentally and financially (Dorfman).

The inability to obtain a comprehensive mailing list of retired agricultural economics faculty from our professional directories has implications for our professional associations. The willingness of retired faculty to spend time on professional activities suggest that professional associations have failed to fully utilize the human capital embodied in many of our retired colleagues. A restructuring of professional association dues and activities to encourage greater participation by retired faculty is worthy of consideration.

The extent to which the findings of this re-

search can be generalized merits discussion. First, the survey population was limited to agricultural economics faculty who had retired from land grant universities. Agricultural economists who had retired from government and private industry were not examined in the survey and, hence, model parameters may not be applicable to these groups. Second, the age of the respondent population enabled many extension faculty to be eligible for the pre-1983 Civil Service Retirement Program. Social Security Amendments of 1983 mandated Social Security coverage for all new federal employees hired after 1983 (Cowen et al.). This provision led to the creation of the Federal Employees' Retirement System (FERS). With retirement benefits under FERS coordinated with Social Security coverage, many of the generous retirement provisions previously available to CSRS participants have been eliminated along with incentives for early retirement. Hence, model parameters may not be valid for extension faculty under the revised FERS.

Other changes taking place in the agricultural economics profession may also change individual retirement strategies. Among these are faculty retrenchment, the changing pattern of twelve-month appointments, changes in the mission of agricultural economics departments, and the consolidation of teaching and research programs on a regional setting. Changes which result in a general decline in the demand for agricultural economics faculty might advance programs for early retirement and result in a reduction in the average retirement age of agricultural economics faculty. By contrast, the abandonment or liberalization of mandatory retirement laws might delay the retirement decision of some faculty, but their impacts on increasing the average age of retirement are expected to be small.

[Received September 1989; final revision received November 1990.]

References

- Anderson, Kathryn H., and Richard V. Burkhauser. "The Retirement-Health Nexus: A New Measure of an Old Puzzle." *J. Human Resour.* 20(1985):315-29.
- Bazzali, Gloria, J. "The Early Retirement Decision: New Empirical Evidence on the Influence of Health." *J. Human Resour.* 20(1985):214-34.
- Beilock, Richard P., Leo C. Polopolous, and Mario Correal. "Ranking of Agricultural Economics Departments by Citations." *Amer. J. Agr. Econ.* 68(1986):595-604.
- Broder, Josef M., and Rod F. Ziemer. "Assessment of Journals Used by Agricultural Economists at Land Grant Universities." *S. J. Agr. Econ.* 16(1985):167-172.
- . "Determinants of Agricultural Economics Faculty Salaries." *Amer. J. Agr. Econ.* 64(1982):301-3.
- Brorsen, B. Wade. "Observations on the Journal Publication Process." *N. Cent. J. Agr. Econ.* 9(1987):315-21.
- Cowen, Jamie, Peggy Crenshaw, Peter R. Lynn, and Joe Richardson. *Your Federal Employees' Retirement System (FERS)*. Alexandria VA: Government Retirement and Benefits, Oct. 1986.
- Dorfman, Lorraine T. "Retired Academics and Professional Activity: A British-American Comparison." *Res. in Higher Educ.* 22(1985):273-89.
- Eason, Marc A. "Southern Region Benefit Survey." Mimeographed. Athens: Agr. Bus. Office, University of Georgia, April 1989.
- Erven, Bernard L. "Reforming Curricula: Challenge and Change for Agricultural Economists." *Amer. J. Agr. Econ.* 69(1987):1037-42.
- Fields, Gary S., and Olivia S. Mitchell. "Economic Determinants of the Optimal Retirement Age: An Empirical Investigation." *J. Human Resour.* 19(1984):245-62.
- Honig, Marjorie, and Giora Hanoch. "Partial Retirement as a Separate Mode of Retirement Behavior." *J. Human Resour.* 20(1985):21-46.
- Knight, Thomas O., S. R. Johnson, and Robert M. Finley. "Extension Program Evaluation Using Normative Decision Models." *Amer. J. Agr. Econ.* 69(1987):338-48.
- Leigh, Duane E. "Why Is There Mandatory Retirement? An Empirical Reexamination." *J. Human Resour.* 19(1984):512-31.
- Litzenberg, Kerry K., and Vernon E. Schneider. "Competencies and Qualities of Agricultural Economics Graduates Sought by Agribusiness Employers." *Amer. J. Agr. Econ.* 69(1987):1031-36.
- McDowell, George R. "The Political Economy of Extension Program Design: Institutional Maintenance Issues in the Organization and Development of Extension Programs." *Amer. J. Agr. Econ.* 67(1985):717-25.
- Pope, Rulon D., and Arne Hallam. "A Confusion of Agricultural Economists—A Professional Interest Survey and Essay." *Amer. J. Agr. Econ.* 68(1987):338-48.
- U.S. Department of Health and Human Services. *Retirement*. Social Security Administration Pub. No. 05-10035, Washington DC, Jan. 1990.
- U.S. Office of Personnel Management. *Your Retirement System*. Pamphlet No. 18, Washington DC, April 1982.

Local Economic Conditions and Wage Labor Decisions of Farm and Rural Nonfarm Couples

J. G. Tokle and Wallace E. Huffman

Effects of geographical differences in local economic conditions on wage labor demand and wage labor participation decisions of rural couples are examined for Current Population Survey households 1978–82. Wage premiums are shown to exist for localities anticipating labor demand growth, higher unemployment rates, larger share of employment in services, and higher costs of living. These effects are stronger for males than females. Effects of local economic conditions on the probability of wage work are consistent with expected market wage and reservation wage effects, and for farm households the probability of wage work increases when expected farm output prices decline or the wage increases.

Key words: farm couples, human capital, local business cycles, local labor markets, rural, wage labor participation, wage rates.

The objective of this paper is to examine the effects of geographical differences in local economic conditions on wage labor demand and labor force participation decisions of U.S. farm and rural nonfarm couples. One hypothesis is that localities with higher anticipated employment growth and unemployment rates pay a wage premium to attract workers from other localities. Thus, when the local anticipated employment growth rate declines (relative to the national rate) real wage rates fall in these localities and rise in others. Another hypothesis is that localities experiencing unanticipated negative labor market shocks show a decline of real wage rates; an example is the case of the upper Midwest during 1981–82. Other geographical wage differences are the result of cost of living and locational amenities.

A considerable amount of research on labor supply of married males and females has utilized a single-worker model (DaVanzo, DeTray, and Greenberg; Mroz, table 1). Furthermore, Mroz has continued this tradition in a

recent analysis of the sensitivity of female labor supply to an array of economic and statistical assumptions. However, we believe that new insights can be gained by considering labor supply decisions in a two-worker, husband-wife model, e.g., see Huffman and Lange. In particular, a spouse's wage has substitution and income effects rather than just income effects alone. A spouse's education may affect a partner's labor supply decision by changing the reservation wage through efficiency effects on family wealth and household production (and for farm households on farm production) or taste. A couple's labor supply decisions are affected by the same economic shocks, so a joint estimation procedure should increase the statistical efficiency of the parameter estimates.

The paper has the following organization: The empirical setting for the study is presented first. Second, the economic models of labor demand and labor supply are developed, including local labor market effects. Third, the data and econometric model are summarized. Fourth, the econometric results are reported. Finally, some conclusions and implications are presented.

The Empirical Setting

This section presents summary measures of differences across localities in labor market char-

The authors are, respectively, an assistant professor, Idaho State University, and a professor, Iowa State University.

Journal Paper No. J-14190 of the Iowa Agriculture and Home Economics Experiment Station Project No. 2738.

Financial assistance was provided by a grant from the Ford Foundation through the Aspen Institute and by the Iowa Agriculture and Home Economics Experiment Station.

Helpful comments were obtained from two anonymous referees, Peter Orazem, and Helen Jensen. Perry Warjiyo provided valuable research assistance in refitting the models.

acteristics, primarily unemployment rates, employment growth, and shocks to labor demand. The basic geographic unit is a state because states are the smallest unit for which annual data are available. Statistics are reported for the twenty-three states having the largest rural population in 1980.

Equilibrium unemployment rates differ geographically and have not tended to converge over time. Table 1 presents the ratio of predicted state unemployment rates to the predicted national unemployment rate for 1970, 1974, 1978, and 1982.¹ Only four years are reported for each state to conserve space; these years lead up to and include the 1978–82 period covered in this study. At any point in time important regional differ-

ences in unemployment rates exist and tend to persist over time (e.g., Hall, Abowd and Ashenfelter, Adams, Topel). For example, the unemployment rates are uniformly higher than average for Michigan, West Virginia, Mississippi, and Louisiana. The unemployment rates are uniformly below average for Pennsylvania, Minnesota, Iowa, Virginia, North Carolina, Georgia, and Texas. For other states, no simple relationship exists during 1970–82 between their unemployment rate and the national average rate.

Table 1 also presents the ratio of predicted state employment growth rates to the predicted national employment growth rate.² The table shows that the employment growth rates in Georgia, Florida, Louisiana, Texas, and Cali-

¹ Each state's annual unemployment rate, 1968–82, and the national unemployment rate were regressed on quadratic trend. The predicted unemployment rates are forecasts from these regressions.

² The natural logarithm of total private sector employment, 1968–82, was regressed on quadratic trend. Predicted employment growth rates are the first differences of the predicted employment values from these regressions.

Table 1. Geographical Distribution of Unemployment and Employment, Twenty-three Selected States, 1970–82

States	Unemployment ^a				Employment ^b				Rural population 1980 (1,000)
	1970	1974	1978	1982	1970	1974	1978	1982	
New York	1.09	1.25	1.17	.94	-.32	-.04	.22	.46	2,700
Pennsylvania	.89	.98	1.08	.90	.45	.35	.26	.17	3,643
Ohio	.89	.92	1.07	1.27	.86	.58	.30	.25	2,879
Indiana	.87	.85	1.04	1.33	1.09	.71	.30	-.04	1,965
Illinois	.83	.82	.97	1.20	.41	.40	.39	.38	1,908
Michigan	1.28	1.26	1.42	1.63	1.09	.66	.26	-.13	2,711
Wisconsin	.85	.74	.85	1.10	1.23	1.02	.82	.63	1,685
Minnesota	.83	.74	.75	.78	1.41	1.28	1.12	.97	1,351
Iowa	.63	.51	.64	.94	1.64	1.06	.52	-.00	1,206
Missouri	.87	.79	.86	1.00	.77	.71	.65	.59	1,567
Virginia	.72	.74	.78	.80	2.09	1.68	1.29	.92	1,817
W. Virginia	1.35	1.05	1.15	1.44	1.32	.97	.60	.29	1,244
N. Carolina	.80	.80	.85	.90	1.59	1.33	1.08	.84	3,059
S. Carolina	1.02	.92	.97	1.11	2.18	1.64	1.08	.55	1,433
Georgia	.83	.92	.92	.84	1.59	1.42	1.29	1.13	2,054
Florida	.85	1.08	1.06	.83	2.45	2.17	1.94	1.72	1,533
Kentucky	.96	.79	.89	1.15	2.09	1.42	.73	.13	1,799
Tennessee	.87	.80	.96	1.23	1.82	1.28	.78	.28	1,818
Oklahoma	.98	.87	1.10	1.52	1.73	1.33	.95	.59	1,556
Mississippi	2.11	1.57	1.25	1.00	2.09	1.64	1.08	.55	1,328
Louisiana	1.30	1.11	1.06	1.05	1.32	1.55	1.77	2.02	1,319
Texas	.78	.75	.74	.70	1.68	1.90	2.16	2.39	2,896
California	1.35	1.28	1.15	.93	1.23	1.37	1.51	1.60	2,060
United States	4.6	6.1	7.2	8.1	2.20	2.26	2.32	2.38	59,495

^a Numbers are ratios of the predicted state unemployment rate to the predicted national unemployment rate. Annual unemployment rates for each state and for the U.S. were regressed on quadratic trend, 1968–82, to obtain the predictions.

^b Numbers are ratios of the predicted state employment growth rate to the predicted national employment growth rate. The logarithm of annual employment in the private sector for each state and the U.S. were regressed on quadratic trend, 1968–82. The predicted employment growth is the difference between the predicted logarithm of employment in two adjacent years.

ifornia are uniformly larger than average, and for New York, Pennsylvania, Ohio, Illinois, and Missouri, employment growth is uniformly below the national average. During this period, some states (i.e., Iowa, West Virginia, South Carolina, Kentucky, Tennessee, Alabama, and Mississippi) went from having far above- to far below-average employment growth rates. In general, these data show a very unequal geographical distribution of employment growth 1970–82.

The geographical distribution of labor demand “shocks” is also unequal. Table 2 presents indexes of local labor demand shocks for the subperiod 1978–82. These measures were constructed as follows. First, the natural logarithm of annual private employment (1968–82) was regressed on quadratic trend. The residuals from these regressions, ϵ_t^s , are indexes of time-varying local demand conditions in state s and year t . Second, the natural logarithm of the national aggregate employment was regressed on quadratic trend. The residuals from this regression, ϵ_t , measure the aggregate labor demand

disturbance in year t . Relative local labor demand disturbances in year t and state s are then defined as $\eta_t^s = \epsilon_t^s - \epsilon_t$, which expresses the current local labor demand “shock” as a deviation from the aggregate labor demand “shock.” (This measure is used by Topel, p. S129).

These measures of local labor demand disturbances have two important features. First, for a given year, substantial geographical variation in employment disturbances occurs. They also are largely unrelated to the aggregate business cycle. Second, for a given state the successive disturbances occur as local cycles. Shocks with the same sign tend to persist for a couple of years then to reverse themselves.

When workers and firms are immobile, local economic conditions will affect real wage rates. Although workers are largely immobile in the short run, they are geographically mobile in the long run. However, differences among localities in their net advantages to firms (industries) and in their amenities and cost of living characteristics to workers contribute to permanent differences in labor markets of different localities.

Table 2. Relative Employment Disturbance, Twenty-three Selected States, 1978–82

States	1978	1979	1980	1981	1982
New York	-2.25	-2.24	-0.97	0.69	2.94
Pennsylvania	-0.40	-0.35	-0.31	0.25	0.32
Ohio	0.67	0.98	-0.34	-0.13	-0.31
Indiana	2.58	2.22	-1.21	-0.49	-1.28
Illinois	0.90	0.71	0.81	-1.05	-0.90
Michigan	4.12	3.05	-0.57	-1.25	-2.16
Wisconsin	0.60	1.42	0.33	-0.62	-0.77
Minnesota	0.66	1.48	1.00	-0.28	-1.56
Iowa	1.89	0.95	0.11	-0.53	-1.87
Missouri	1.22	1.46	-0.37	-0.98	-0.37
Virginia	-0.52	-1.41	0.49	-0.15	0.34
W. Virginia	0.41	1.98	0.62	-1.40	-1.23
N. Carolina	-0.17	0.42	0.12	0.03	-0.17
S. Carolina	-0.72	-0.67	0.06	0.57	0.47
Georgia	-0.43	-0.80	-0.49	0.11	1.65
Florida	-3.33	-2.85	0.14	1.88	2.55
Kentucky	1.01	1.29	-0.86	-1.13	-0.01
Tennessee	0.87	0.44	-0.77	0.39	-0.18
Oklahoma	-1.50	-2.56	-1.15	0.75	2.00
Mississippi	1.95	1.49	0.26	-0.97	-1.34
Louisiana	0.92	-1.03	0.23	0.40	-1.29
Texas	-1.09	-1.42	-0.72	0.76	0.68
California	0.67	0.85	0.84	0.12	-1.31
United States	2.39	3.59	1.88	0.34	-3.77

Note: The numbers are measures of $\eta_t^s = \epsilon_t^s - \epsilon_t$, except for the last row. See the text for details.

The Economic Model

In a labor market workers simultaneously sell their services and buy the attributes of jobs, including location. Employers also buy the services and characteristics of workers and sell job attributes. Therefore, the theory of equalizing wage differentials stresses both supply and demand for labor, and the market equilibrium process allocates or assigns specific workers to specific firms and locations (Rosen).

Labor Demand

The actual wage can be viewed as the summation of two distinct transactions, one for labor services and worker characteristics and another for job and locational attributes (Rosen). Localities are assumed heterogeneous over characteristics that matter for optimal firm-industry (employer) and household (worker) decisions. This environment contains spatially related competitive labor markets with equilibrium wage differentials. A wage premium is paid by employers to induce workers to undertake undesirable tasks or work at an undesirable location. Unequal transitory shocks to labor demand in different localities are an added source of geographical differences in wage rates.

The wage elasticity of aggregate labor demand for a locality is negative, but individuals face a perfectly elastic demand for their labor. Our labor demand or wage rates at a given location and for a given sex are assumed to depend on skill or human capital (ζ) and job or locational characteristics—permanent or anticipated labor market conditions (Ω), transitory or unanticipated labor market conditions (ω), local cost of living (ψ), and locational amenities (Δ). This relationship is summarized as

$$(1) \quad W^j = W^j(\zeta^j, \Omega, \omega, \psi, \Delta),$$

where $j = M(\text{males}), F(\text{females})$.

Permanent or anticipated labor market conditions refer to important conditions that local firms and workers know and include in rational decision making. These variables include the anticipated or long-run unemployment rate, anticipated rate of job growth, and anticipated change in the occupational composition of employment. When anticipated unemployment harms workers (households) more than firms (employers), firms or localities with higher anticipated unemployment rates will have higher

equilibrium wage rates than other firms or localities (Hall, Abowd and Ashenfelter, Adams, Topel). That is, higher wage rates compensate for higher probabilities of unemployment. When households (workers) bear most of the cost of geographical and occupational mobility, geographic differences in the anticipated rate of employment growth and in change of occupational composition of employment are also the source of compensating geographical wage differentials (Lilien, Adams, Topel). Thus, positive wage differences provide economic returns to households for geographical or occupational mobility.

Locational differences in the cost of living can be decomposed into effects of prices of goods traded among locations and of nontraded goods and services (Tolley, Kenny and Denslow). With competitive markets, the prices of traded goods in two areas differ by transport costs so that they likely are an insignificant source of cost of living differences. However, prices of nontraded goods and services differ more between two localities. The price of housing-plus-access is a good example for the nontraded good. Furthermore, because housing costs are about 14% to 15% of workers' household expenditures, locational differences in these prices are a significant source of interlocality wage differences.

Areas differ in the quantity and quality of locational amenities. Normal climatic conditions, e.g., average January and July temperature, are characteristics of the local environment that likely matter to households (workers) and possibly to firms. Other studies (e.g., Israeli, Hock and Drake, Kenny and Denslow) have found significant effects of climatic conditions on wage rates.

Labor Supply and Labor Force Participation

Households make labor supply and labor force participation decisions of their members. In this analysis, the focus is on single-family, husband-wife households, where the decisions are for one period, and households are risk neutral about uncertain outcomes including employment or unemployment. This means the expected wage is used in resource allocation decisions.

Participation decisions of farm and rural nonfarm households are analyzed. The nonfarm households are assumed to have only wage and asset income; farm households have asset income, self-employment income from their farm business, and perhaps off-farm wage income. Thus, the wage-labor participation decisions of

farm households are more complex than for rural nonfarm households (Huffman and Lange, Strauss).

The economic decision-making framework of these households is summarized in equations (2)–(5).

$$(2) \quad U = U(T_h^M, T_h^F, Y; \zeta^M, \zeta^F, \Delta, \tau)$$

$$(3n) \quad \bar{T} = T_m^j + T_h^j, T_m^j \geq 0, j = M, F$$

$$(3a) \quad \bar{T} = T_f^j + T_m^j + T_h^j, T_m^j \geq 0, j = M, F$$

$$(4n) \quad (1 - u^M) W_0^M T_m^M + (1 - u^F) W_0^F T_m^F + V = P_y Y$$

$$(4a) \quad (1 - u^M) W_0^M T_m^M + (1 - u^F) W_0^F T_m^F + V + P_Q Q - W_x X + V = P_y Y$$

$$(5a) \quad Q = Q(T_f^M, T_f^F, X; \zeta^M, \zeta^F, \Delta).$$

Notation “*n*” and “*a*” after the equation number refer to nonfarm and farm households, respectively; other equations refer to both types of households.

Farm and nonfarm households derive utility from the leisure time of the husband and wife (T_h^j) and from goods purchased in the market (Y). Household utility also depends on husband's and wife's human capital (ζ^j), local climate (Δ), and other household characteristics (τ), e.g., number of children in the household and commuting distance to service centers.

specific unemployment rates are exogenously determined for the households. Furthermore, firms and households do not know the unanticipated parts of employment growth or unemployment rates. W_o^j is the anticipated wage, given unemployment, when unanticipated local labor disturbances are zero, i.e., $\omega = 0$ in equation (1). Expected household wage income is $(1 - u^M) W_o^M T_m^M + (1 - u^F) W_o^F T_m^F$. Household asset income is V . Farm households also have uncertain self-employment or net income from a farm business ($P_Q Q - W_x X$). Farm output (Q) is produced by inputs of husband's and wife's farm hours (T_f^j) and by purchased inputs (X). The efficiency of the production process is affected by human capital of the husband and wife (ζ^j) and climate (ϕ).³

Households are assumed to face a perfectly elastic supply of the consumption good (Y) at a price P_y . Farm households also face a perfectly elastic supply of inputs at price W_x and perfectly elastic demand for farm output. However, the price of farm output is uncertain when production plans are made, and P_Q denotes the expected price.

For rural nonfarm households, wage labor supply functions are obtained by maximizing (2) subject to (3n) and (4n). For farm households, wage labor supply functions are obtained by maximizing (2) subject to (3a), (4a), and (5a). These wage labor supply equations are:

$$(6n) \quad T_m^j = S_m^j[(1 - u_o^j)W_o^j, (1 - u^k)W_o^k, P_y, V, \zeta^M, \zeta^F, \Delta, \tau], T_m^k > 0, \\ S_m^j[(1 - u^j)W_o^j, P_y, V, \zeta^M, \zeta^F, \Delta, \tau], T_m^k = 0, \quad j, k = M, F; j \neq k.$$

$$(6a) \quad T_m^j = S_m^j[(1 - u^j)W_o^j, (1 - u^k)W_o^k, P_y, V, P_Q, W_x \zeta^M, \zeta^F, \Delta, \tau, \phi], T_m^k > 0; \\ S_m^j[(1 - u^k)W_o^j, P_x, V, P_Q, W_x, \zeta^M, \zeta^F, \Delta, \tau, \phi], T_m^k = 0, \quad j, k = M, F; j \neq k.$$

Farm and nonfarm households receive human time endowments each year (\bar{T}) for the husband and wife, which are considered heterogenous. In nonfarm households, the time of each adult is allocated between work for a wage (T_m^j) and leisure (T_h^j). In farm households, time is allocated among work on their own farm (T_f^j), work for a wage (off farm) (T_m^j), and leisure (T_h^j). In farm and nonfarm households, optimal hours of wage work might be zero in any year. Hence, a nonnegativity constraint is imposed on wage work ($T_m^j \geq 0$).

The cash income of farm and rural nonfarm households is uncertain because of uncertain employment prospects. Thus, the wage rate, given employment, is adjusted for the expected probability of unemployment (u^j). These sex-

Equations (6n) and (6a) show that wage labor supply functions for a given individual have different structures (Huffman and Lange). When both married nonfarm individuals work for a wage, their wage labor supply is a function of their expected anticipated wage rates, price of consumption goods, asset income, their human capital stocks, the local climate, and tastes. For farm households, the expected price of farm output, the price of purchased farm inputs and a technology parameter also determine off-farm wage labor supply.

Wage-labor participation decisions of an in-

³ The effect of a tax on income is excluded from the economic (and econometric) model. This simplification is unlikely to be of major consequence for the empirical results (Mroz).

dividual are modeled as a comparison of his (her) reservation wage and anticipated market wage. The reservation wage equation is derived from the wage-labor supply equation by setting wage work hours equal to zero:

$$(7n) \quad W_r^j = \left(\frac{1}{1 - u^j} \right) G_r^j [P_y, V, \zeta^M, \zeta^F, \Delta, \tau, (1 - u^k), \Omega, \omega = 0, \psi], \\ j, k = M, F; j \neq k.$$

$$(7a) \quad W_r^j = \left(\frac{1}{1 - u^j} \right) G_r^j [P_y, V, P_Q, W_X, \zeta^M, \zeta^F, \Delta, \tau, \phi, \\ (1 - u^k), \Omega, \omega = 0, \psi], \\ j, k = M, F; j \neq k.$$

A nonfarm household member participates in wage work when his (her) reservation wage is less than the anticipated wage in the market.⁴ A farm household member participates in nonfarm wage work when the marginal value of his (her) leisure and (or) farm work hours are less than the anticipated nonfarm wage.

The probability of wage work can be expressed as the probability that an individual's reservation wage is less than his (her) anticipated market wage. For the i th household and j th married individual, define

$$D_i^j = \begin{cases} 1 & \text{if } j\text{th individual works for a wage} \\ 0 & \text{otherwise} \end{cases}, \\ j = M, F;$$

then the probability of wage work for the i th individual is

$$(8) \quad P_r\{D_i^j = 1\} = F\{W_{ri}^j < W_{oi}^j\}; j = M, F, \text{ or}$$

$$(8a) \quad P_r\{D_i^j = 1\} = F[P_y, V, \zeta^M, \zeta^F, \Delta, \tau, \Omega, \\ (1 - u^j), (1 - u^k)], \\ j, k = M, F; j \neq k, \text{ and}$$

$$(8n) \quad P_r\{D = 1\} = F[P_y, V, P_Q, W_X, \phi, \zeta^M, \zeta^F, \Delta, \tau, \Omega, \\ (1 - u^j), (1 - u^k)], \\ j, k = M, F; j \neq k,$$

where $F(\cdot)$ is a distribution function. Variables that explain the probability of wage work enter an individual's labor demand and labor supply functions, except for the individual's anticipated

wage rate. When the labor supply schedule has a positive slope, variables that cause the labor supply curve to shift to the left will increase the reservation wage and reduce the probability of wage work. A change in a variable that raises the market wage—raises the labor demand curve—will increase the probability of wage work.

Selected variables are examined for their effects on the probability of wage work. An increase of an individual's schooling will increase the anticipated market wage and reservation wage. The net effect on the probability of wage work is a priori ambiguous, but other studies (e.g., Heckman and MaCurdy 1980, 1982; Huffman and Lange) have found a strong positive effect of an individual's schooling on the probability of wage work for married farm and nonfarm males and females in the United States.

When leisure is a normal good, a higher expected price of farm output reduces the probability of wage work by couples who operate a farm business. The reason is that the quantity demanded of husband's and wife's farm labor increases, provided their farm hours are a normal input and the quantity of their leisure demanded increases due to the increase of expected profit. Higher farm input prices change the probability of wage work in an a priori ambiguous direction. Hired (nonfamily) labor and family labor are heterogenous. If hired farm labor and husband's and wife's farm labor are gross substitutes, then a higher wage for hired labor will cause a rightward shift of the demand for husband's and wife's farm labor. A higher input price reduces expected farm profit and reduces the demand for leisure of the husband and wife. Thus, the net effect on the reservation wage and probability of wage work is a priori ambiguous. A change in the price of other inputs yields a similar conclusion.

The effects of local labor market conditions depend partially upon expectations. If the demand for leisure increases as the expected wage decreases, a higher expected local unemployment rate will raise an individuals' reservation wage. When anticipated unemployment hurts workers more than firms, firms or localities that have higher expected unemployment rates will pay higher wage rates. Thus, the net effect of anticipated unemployment on the probability of wage work depends on which of these changes is largest. Localities having more rapid anticipated job growth are expected to pay higher wage rates than other areas and hence will increase the probability of wage work. On the other hand,

⁴ Alternatively, decisions can be made by comparing indirect utility functions associated with different outcomes.

unanticipated changes in local labor market conditions are not expected to affect the probability of wage work. These disturbances can at best be poorly forecasted and at worst cannot be forecasted. The anticipated wage rates are higher in these localities, while the reservation wage likely is unaffected.

The Data and Econometric Model

The econometric model incorporates effects of local economic conditions on wage labor demand and probability of wage work of married rural males and females. Focusing on males and females provides a better picture of the total effect on households and on differences resulting from gender. Farm and rural nonfarm households from the Current Population Surveys (CPS) and having economic activity in 1978, 1979, 1981, and 1982 are the decision units to be studied.

The Data

The CPS is a monthly survey containing information about the employment status of members of approximately 60,000 interviewed households residing in every state of the United States. The annual demographic file (March) contains information on labor force participation, employment, and earnings of household members during the calendar year before the survey. Starting in 1977, the state of residence and farm-nonfarm residence of each household are identified.

One sample consists of nonmetropolitan-nonfarm households where the husband and wife are present and no self-employment income was received. This is the closest definition to rural nonfarm because households having rural nonfarm residence are not identified in the CPS. This sample comprises "rural nonfarm" wage earning households. The other sample consists of husband-wife households that have a farm residence and self-employed income from farming. In both samples, households having a residence in the contiguous forty-eight states, except for New England, are included. Households in New England were excluded because of the lack of importance of agriculture in those states. These two samples from the forty-two states consist of about 8,115 rural nonfarm households and 1,466 farm households per year.

Local markets are defined as state units in this

study. A state is the smallest geographic unit in which a CPS household can be identified. States are also the smallest political-economic-geographic unit for which annual data are collected on employment, unemployment, and agricultural prices. In addition, government programs frequently target state units. Other recent studies (e.g., Adams, Topel) have used state units as labor markets.

The years 1978–82 give variations in local economic conditions that affect wage and participation rates. This period was chosen for several reasons. The CPS first identified the farm-nonfarm residence of households in 1977. Second, the period 1975–79 is the trough-to-peak part of a national business-cycle expansion (Executive Office of the President). The national average unemployment declined from 8.3% in 1975, to 5.8% in 1979. Net farm income was relatively good in the late 1970s. The period starting in 1980 contained a business-cycle contraction. The national unemployment rate rose from 7% in 1980 to 9.5% in 1982–83. The sharp rise of real interest rates and fall in the value of the U.S. dollar contributed to the drop in net farm income during 1981 and 1982. Although the depression of the farm economy continued after 1982, extending the analysis through 1983 did not seem wise because the first large government payment-in-kind (PIK) programs occurred in 1983. Third, the data in table 2 show relatively large geographical variation in shocks to labor demand during this period. Fourth, Lilien has shown that employment growth was unequally distributed across U.S. industries in the 1970s. Significant shifts of the occupational-industrial mix of employment occurred, especially a rise in the share of employment in services, finance, insurance and real estate but a decrease in manufacturing.

When the four cross-sectional files are combined, the data on forty-two states and four years gives 168 potentially distinct observations on different local economic conditions. Although data for four years are used, it is not a panel consisting of the same households. About 25% of the CPS households in any year are replaced.

Empirical Definitions

Definitions and sample mean values of the variables are presented in table 3. More details are presented below on the derivation of selected variables.

Average hourly wage rates are derived for ru-

Table 3. Variable Names and Sample Means, Rural Married Couples 1978–79, 1981–82

Symbol	Variable Description	Sample Mean	
		Nonfarm	Farm
<i>Individual/household</i>			
AGEM	Husband's age (yrs)	47.0	50.5
AGEF	Wife's age (yrs)	43.9	47.2
EDM	Husband's schooling (yrs)	11.5	11.3
EDF	Wife's schooling (yrs)	11.6	11.8
RACE	1 if nonwhite; 0 otherwise	.07	.03
KIDS06	Number of children under age 6	.31	.27
KIDS618	Number of children ages 6–18	.66	.69
ASSETINC	Household real nonwage and nonfarm income (1967 prices) ^a	\$10,386 ^a	\$10,701 ^a
<i>Local labor market conditions</i>			
PJOBGR	Predicted state employment growth rate (see text)	2.17	2.06
PURATE	Predicted state unemployment rate (see text)	7.46	7.18
ΔSHRSER	Change of share of a state's jobs in serv. occupation (previous 2 yrs.)	1.02	.91
ESHOCK	Relative state employment growth shock (see text)	.08	.15
RURATE	Residual state unemployment rate (see text)	−.37	−.38
<i>Cost of living and locational amenities</i>			
PLAND	State average price of agricultural land in 1978 (\$1,000/acres)	.91 ^a	.79 ^a
URBAN	Percentage of state population urban	.68	.66
JAN	Normal January ave. temperature (degree F)	34.1	29.7
JULY	Normal July ave. temperature (degree F)	75.9	75.1
<i>Agricultural prices and climate</i>			
PCROP	State real price index for crops (1967 prices)		.444 ^a
PLIVE	State real price index for livestock (1967 prices)		.517 ^a
FARMWAGE	State real wage rate for hired farm labor (1967 prices)		.504 ^a
POTINP	State real price index for nonlabor farm input (1967 prices)		.536 ^a
RAIN	State annual average precipitation (inches)		35.7
GDD	State average growing season length—growing degree days (1,0005)		3.336
<i>Regional dummies and trend</i>			
NC	1 for resident in North Central Region; 0 otherwise	.28	.46
SOUTH	1 for resident in South; 0 otherwise	.51	.38
WEST	1 for resident in West; 0 otherwise	.06	.12
TIME	Trend	3.00	3.00
<i>Dependent variables</i>			
WAGEM	Real married male nonfarm wage (\$/hr, 1967 prices)	\$2.87 ^a	
WAGEF	Real married female nonfarm wage (\$/hr, 1967 prices)	\$1.66 ^a	
D ^M	1 if husband works for wage; 0 otherwise	.75	.43
D ^F	1 if wife works for a wage; 0 otherwise	.54	.39

^a Geometric mean. All other numbers are arithmetic means.

ral nonfarm married males and females. For individuals in the rural nonfarm wage work households, the average wage of an individual is his (her) wages and salaries for the year preceding the survey divided by the product of his (her) hours worked per week last year and weeks worked last year. An accurate measure of the average wage rate cannot be computed for individuals in a household with self-employment income because hours worked include hours at all jobs, both wage work and self-employment. Thus, average wage rates are not available for CPS farm household members. Nominal wage rates are deflated by the consumer price index.

Five measures of local labor market conditions are derived. The predicted state employment growth rate ($PJOBGR_t$) is the difference in forecasted values of the natural logarithm of a state's private sector employment in t and $t - 1$. The forecasts were obtained from a regression of the natural logarithm of employment, 1968–82, on a quadratic trend. A state's unemployment rate was measured for all private sector employees rather than having separate rates for males and females. The predicted state unemployment rate ($PURATE_t$) measures the anticipated local unemployment rate. It is obtained from a regression of a state's annual unemployment rate, 1968–82, on quadratic trend. The change in the share of a state's employment that is in the service sector ($\Delta SHRSER_t$) indicates changes in the occupational mix of local labor demand. It is defined as the share in t minus the share in $t - 2$. Service jobs include employment in services, transportation, government, finance, and wholesale and retail trade. The derivation of the relative shocks to labor demand in state labor markets ($ESHOCK_t$) is described above. In this study, these shocks are unanticipated by firms and workers. The residual or unanticipated unemployment rates ($RURATE_t$) is the actual unemployment rate in t minus the predicted unemployment rate for t .

Geographical differences in the cost of living and locational amenities are tied to cost of housing-plus-access and climate. The price of land is a major part of the cost of housing-plus-access. For households living in rural areas, the base price of land is represented by the average agricultural land price in 1978 (U.S. Dep. of Commerce 1980). In addition, the price of housing-plus-access increases as the percentage of the population living in urban areas increases because the cost of land plus commuting is larger in urban areas. When land prices and wage rates are positively associated and housing and hus-

band's and wife's leisure are substitutes, an increase of P_{LAND} will increase the reservation wage. This makes the effect of P_{LAND} on the probability of wage work a priori uncertain. However, if housing and husband's and wife's leisure are complements, their reservation wage rates will decline. The percentage of the population living in urban areas in 1980 (U.S. Dep. of Commerce 1981) is used as a second proxy for the cost of housing-plus-access. For locational amenities, thirty-year (1950–80) normal average January and average July temperatures (Weiss, Whittington, and Teigen) were used. Kenny and Censlow and Hoch and Drake found nonlinear effects of these temperatures on log wage rates in earlier studies.

The profitability of local agriculture is represented by indexes of agricultural prices and agricultural climate. Indexes of crop prices, livestock prices, wage for farm labor, and prices of other inputs are derived and deflated by the consumer price index. Output price indexes for crops and livestock are constructed because the average labor intensity is significantly different for these output groups (Huffman and Evenson, chap. 10). The crop price index is composed of prices of twenty-six different commodities or commodity groups. The livestock price index is composed of seven commodity groups. The expected prices used to derive the Fisher-type output price indexes are primarily one-year lagged prices of the commodities (Huffman and Evenson, chap. 10).

Input prices are split into two groups, farm labor and other inputs. The price of farm labor is the hourly wage paid to employees working for cash wages only. The nonlabor input prices include fertilizer, feed, capital, seed, land, and miscellaneous inputs. Current prices are used to derive Fisher-type price indexes of these inputs (Huffman and Evenson, chap. 10). Both output and input price indexes are deflated by the consumer price index.

Normal annual precipitation and normal growing-degree-days are important climatic variables for agricultural production. Normal annual rainfall is a twenty-five-year average of annual precipitation (Weiss, Whittington, and Teigen). Natural precipitation is the primary source of water for much of U.S. agriculture. However, in low precipitation areas, irrigation is a costly substitute. Accumulated growing degree days, GDD , is a measure of accumulated heat units from a temperature range that is particularly favorable to corn production (U.S. Dep. Agriculture and U.S. Dep. Commerce; U.S. Dep.

Commerce 1971, 1981). Corn is grown in almost every state, but more generally the index is highly correlated with good growing conditions for warm-season crops.

Household asset income reflects earnings from interest, dividends, and rental property deflated by the consumer price index. Thus, all income- and work-conditioned transfers are excluded.

The Econometric Model

The econometric model consists of two labor demand equations and two wage-participation equations. The empirical specification of the labor demand equations is similar for married males and females:

$$(9) \quad \ln WAGE_i^j = \alpha_1^j + \alpha_2^j EXP_i + \alpha_3^j EXP_i^2 + \alpha_4^j ED_i + \alpha_5^j RACE_i + \alpha_6^j PJOBGR_i \\ + \alpha_7^j PURATE_i + \alpha_8^j \Delta SHRSER_i + \alpha_9^j ESHOCK_i + \alpha_{10}^j RURATE_i \\ + \alpha_{11}^j \ln PLAND_i + \alpha_{12}^j URBAN_i + \alpha_{13}^j JAN_i + \alpha_{14}^j JAN_i^2 + \alpha_{15}^j JULY_i \\ + \alpha_{16}^j JULY_i^2 + \alpha_{17}^j NC_i + \alpha_{18}^j SOUTH_i + \alpha_{19}^j WEST_i + \alpha_{20}^j TIME_i + \alpha_{21}^j \lambda_i^j \\ + \epsilon_i^j, i = 1, \dots, n, j = M, F.$$

The natural logarithm of an individual's real wage is expressed as a function of his (her) own human characteristics—experience, experience squared, education, race—and job/local conditions that are potential sources of geographical wage differentials. The last group of variables includes sets of variables for local labor market conditions, cost of living and locational amenities, and regional dummy variables. A time trend and sample selectivity variables (λ_i^j) are also included in each equation.

Equation (9) is quadratic in experience and in January and July temperatures but not in other variables. This choice was largely based on evidence reported in other studies (e.g., Adams, Topel, Kenny and Denslow) and preliminary fits of the wage equations. Experience is defined as an individual's age minus years of schooling completed minus 6. This measure of experience is a reasonable proxy for useful work experience of males but less so for females engaged in household activities. It is less endogenous to current labor market decisions than actual experience (Heckman and MaCurdy). When a sample selection variable is included in the wage equation, sample selection bias is unlikely to be associated with the use of a work experience variable (Mroz).

The regional dummy variables may contain redundant information about labor demand. If the set of variables representing local labor market conditions and cost of living and locational

amenities proxy relatively well the sources of geographical wage differentials, then the regional dummy variables will not make a statistically significant contribution to real wage rates. When women spend more time out of the labor force and are more geographically immobile than men, women's wage rates likely will be less responsive to their measured human capital and to local economic conditions (Mincer). The coefficients of variables in the female wage equation then will be smaller than in the male equation.

Sample selectivity and autocorrelation are potential problems in the wage equation. If a sample selection variable was not included in equation (9), the disturbance term of the wage equation would have a nonzero mean because the equation is fitted to a nonrandom subset of

the total population (Heckman). This is a potential source of statistical bias in estimated coefficients. With a selection variable included, the disturbance has a zero mean and a normal distribution.

The data files consist of two adjacent-year cross sections (1978–79, 1981–82) that are separated by one year. Some autocorrelation might be expected in adjacent year observations on the same individual, but observations that are one or more years apart are less likely to be correlated. This short, disjointed time series, where the composition of the sample changes over time, is not conducive to correction for autocorrelation. Failure to correct for autocorrelation when it is present results in some loss of estimation efficiency, but the least-squares estimator is unbiased or consistent (Johnston). Contemporaneous correlation of disturbances in the two wage equations might occur. However, because the observations in the wage equations of the married males and females are not equal, the two wage equations of the married males and females are not equal, the two wage equations cannot be estimated jointly in the seemingly unrelated regression model. Ignoring contemporaneous cross-equation correlation of disturbances results in some loss of estimation efficiency (Johnston).

The empirical specification of the probability of the j th married individual in the i th farm household participating in wage work is

(10a)

$$\begin{aligned} \Pr(D_i^j = 1) = F[& \beta_1^j + \beta_2^j AGEM_i + \beta_3^j AGEM_i^2 + \beta_4^j EDM_i + \beta_5^j EDF_i + \beta_6^j RACE_i \\ & + \beta_7^j KIDS06_i + \beta_8^j KIDS618_i + \beta_9^j \ln ASSETING_i + \beta_{10}^j PJOBGR_i + \beta_{11}^j PURATE_i \\ & + \beta_{12}^j \Delta SHRSE_i + \beta_{13}^j ESHOCK_i + \beta_{14}^j RURATE_i + \beta_{15}^j \ln PLAND_i + \beta_{16}^j URBAN_i \\ & + \beta_{17}^j JAN_i + \beta_{18}^j JAN_i^2 + \beta_{19}^j JULY_i + \beta_{20}^j JULY_i^2 + \beta_{21}^j \ln PCROP_i \\ & + \beta_{22}^j \ln PLIVE_i + \beta_{23}^j \ln FARMWAGE_i + \beta_{24}^j \ln POTINP_i + \beta_{25}^j GDD_i + \beta_{26}^j RAIN_i \\ & + \beta_{27}^j RAIN_i \bullet GDD_i + \beta_{28}^j NC_i + \beta_{29}^j SOUTH_i + \beta_{30}^j WEST_i + \beta_{31}^j TIME_{ij}] j = M, F, \end{aligned}$$

where $F(\cdot)$ is the normal distribution function. The participation equations for rural nonfarm household members are similar, except that prices of agricultural outputs and inputs, annual precipitation, and the length of growing season are excluded. This probability is a function of a set of variables representing an individual's own characteristics, spouse characteristics, household characteristics, and local conditions.⁵ A time trend is also included.

Equation (10) is a reduced-form specification, including only variables in the labor supply or labor demand functions (Huffman and Lange). The husband's age (and age squared) controls for nonlinear life-cycle and work-experience effects when education is held constant. Age and experience are highly correlated, so experience is not included as a separate variable. Because husband's and wife's ages are also highly correlated, only the husband's age is used for explaining both the husband's and wife's probability of wage work.

Given that participation decisions of a husband and wife in farm and nonfarm households are assumed to be joint within a household-optimizing framework, the probability of a married individual participating in wage work is affected by some of their spouse's characteristics. Also, these decisions are affected by random or unmeasured shocks to labor supply (reservation wage) and labor demand. These shocks likely are across spouses (Huffman and Lange). Thus, the estimation procedure for equation (10) is bivariate probit. If a husband and wife are affected similarly by a given shock, the correlation between these disturbances will be positive, and vice versa.

The Results

Results from testing hypotheses about the effects of local economic conditions and other

variables on labor demand for rural nonfarm married males and females and on the probability of wage labor participation of married males and females in farm and rural nonfarm households are reported in this section.

Labor Demand

Labor demand equations are fitted to the data for 24,571 married rural nonfarm males and 17,508 married rural nonfarm females. These equations (9) are fitted by least squares with an instrumental variable included to control for sample selectivity. Conclusions reached about labor demand for married rural nonfarm males and females likely are applicable to married farm males and females.

Estimates of four labor demand functions are reported in table 4; two for males and two for females. The second equation excludes the regional dummy variables. A test of the null hypothesis that the coefficients of the three regional dummy variables are jointly equal to zero is rejected for males but not for females. The sample value of the F -statistic is 5.08 for males and 1.90 for females, and the critical value of the F -statistic with 3 and infinite degrees of freedom is 3.79 at the 1% significance level. Thus, for males the eleven variables representing local labor market conditions, cost of living differences, and location amenities do not capture all of the geographical differences in wage rates. Furthermore, the long history of wages being lower in the South is no longer supported. Male wages are higher in the South than in other regions, and female wages are not lower.

In the wage equations, all of the coefficients of the human characteristics have expected signs and are significantly different from zero at the 1% level. An increase of an individual's experience has first a positive but diminishing marginal effect on the real wage. The maximum effect occurs at 36.9 years for males and 31.5 years for females. This pattern has been reported in many studies. However, the \log_e wage-experience relationship is more convex for males than

⁵ Although the number of children at home and the amount of household asset income is endogenous in a lifetime planning horizon, Mroz found that these variables were not endogenous in his tests. They are exogenous in our study.

Table 4. Labor Demand Equations: Rural Nonfarm Married Males and Females, 1978–79, 1981–82

Variables	ln Wage			
	Males		Females	
	(1)	(2)	(3)	(4)
Human capital				
<i>EXP</i> (<i>AGE-ED-6</i>)	.031 (20.04)	.031 (19.95)	.017 (12.88)	.017 (12.93)
<i>EXP</i> ² /100	-.042 (10.30)	-.041 (10.20)	-.027 (9.35)	-.027 (9.38)
<i>ED</i>	.055 (40.21)	.055 (40.12)	.071 (30.12)	.071 (30.19)
<i>RACE</i>	-.204 (13.72)	-.203 (13.69)	-.065 (3.48)	-.064 (3.43)
Labor market conditions				
<i>PJOBGR</i>	.016 (2.68)	.024 (4.89)	.009 (1.19)	.011 (1.73)
<i>PURATE</i>	.012 (4.21)	.011 (4.13)	.004 (1.16)	.003 (0.76)
Δ <i>SHRSER</i>	.005 (1.91)	.005 (1.97)	.002 (0.72)	.002 (0.67)
<i>ESHOCK</i>	.005 (1.79)	.004 (1.45)	.005 (1.31)	.004 (1.07)
<i>RURATE</i>	-.006 (1.15)	-.004 (0.74)	-.011 (1.68)	-.010 (1.51)
Cost of living and locational amenities				
ln <i>PLAND</i>	.073 (5.46)	.060 (5.47)	.053 (3.02)	.056 (3.87)
<i>URBAN</i>	.255 (5.77)	.180 (5.33)	.011 (0.20)	.012 (0.26)
<i>JAN</i>	.003 (1.15)	.008 (3.96)	-.002 (0.55)	.002 (0.74)
<i>JAN</i> ² /100	-.001 (3.91)	-.014 (5.29)	.002 (0.57)	-.001 (0.32)
<i>JULY</i>	-.087 (1.43)	-.059 (1.25)	.197 (2.41)	.166 (2.59)
<i>JULY</i> ² /100	.057 (1.41)	.036 (1.14)	-.136 (2.49)	-.117 (2.77)
Regional dummies and trend				
<i>NC</i>	-.023 (1.37)		-.038 (1.67)	
<i>SOUTH</i>	.057 (2.65)		-.015 (0.51)	
<i>WEST</i>	.064 (2.09)		-.002 (0.05)	
<i>TIME</i>	-.026 (6.29)	-.027 (6.42)	-.012 (2.22)	-.013 (2.27)
$\hat{\lambda}$.279 (5.53)	.286 (5.68)	-.020 (0.79)	-.021 (0.82)
Intercept	3.175 (1.41)	2.260 (1.27)	-7.625 (2.50)	-6.404 (2.66)
<i>R</i> ²	.1619	.1614	.0781	.0780
<i>N</i>	24,571	24,571	17,508	17,508

Note: The *t*-ratios are conditioned on the sample selection variables.

for females. This result is consistent with married females having on average less actual labor market experience for any given measured experience than for married males (Mincer and Ofek). Males may also make larger investments in experience during their early work-life than females.

A one-year increase in schooling causes a larger percentage increase of the female than the male wage, 7.1% versus 5.5%. These relative magnitudes are consistent with some results reported in other studies, e.g., Topel for males and Gerner and Zick for females. Although the average wage for married rural nonfarm working females in the CPS sample is 57.8% of the wage for males, the marginal increase of the real wage is larger for males than for females (\$.16 versus \$.12 per hour).

Nonwhite rural nonfarm males earn 20% less than rural nonfarm white males, other measured variables equal, and nonwhite females earn 6.5% less than white females. Topel found an 18% difference in wages of white and nonwhite males for 1976–79. Other studies have also shown large gaps in the wages of white and black males on average but little or no gap in white and black women's wage ratios (e.g., Hammermesh and Rees).

All of the signs of coefficients on variables representing local labor market conditions are positive, except for the unanticipated unemployment rate, in both the male and female wage equations. They are all consistent with hypotheses developed earlier. Most coefficients in the male wage equation are significantly different from zero at the 5% level. For females, the effects of the local labor market variables are smaller for anticipated variables than for males but are as large or larger for unanticipated variables. Except for the unanticipated unemployment rate, the local labor market variables are statistically weaker in the female than in the male wage equations.

The real wage rates of married males incorporate compensation for local market conditions that can be anticipated by employers and workers. The evidence, however, is weaker for females. This outcome is expected when married females spend a smaller share of their time in the labor force and are tied to their husband's locational choice (Mincer). For married males, wage rates are higher in localities with higher expected rates of employment growth and higher expected unemployment rates.

The wage premium in localities having higher expected growth of labor demand compensates

males for the costs of geographical (and possibly occupational) mobility. Topel found a similar effect. Our results show that a 1% increase of the expected growth rate of local employment increases the male real wage by 1.5% to 2.5%. For females, the marginal effect is about 1%. These results imply that married rural males and females in the upper Midwest experienced significant reductions in real wage rates during the early 1980s when employment growth rates fell far below (about 2 percentage points) the national average (see table 1).

Localities with higher anticipated unemployment rates also pay higher rural wage rates for males. This result is similar to the findings of Abov and Ashenfelter, of Adams, and of Topel. Localities (firms) with higher expected unemployment rates pay a premium to compensate male workers for bearing this risk; the evidence is much weaker for females. Again, this finding is consistent with greater immobility of married females than males. These results imply the wage premium associated with higher anticipated unemployment rates is large enough for males to keep expected real wage rates approximately unaffected.

An increase locally of the share of service jobs increases the real wage. The effect is significantly different from zero at the 5% level for males but not for females. Although service jobs span a wide range of skills from motel and restaurant staff to investment bankers, compensation is needed to create the incentives for males to invest in skills and change occupations.

Real wage rates respond to unanticipated changes in local labor market conditions in a way that is consistent with employers and employees sharing good and bad outcomes. Topel found similar effects for wage rates of nonfarm males. The coefficients of *ESHOCK* are significantly different from zero at the 7% level [table 4, equation (1)] in the male wage equation and at the 20% level in the females' wage equation. The coefficient of *RURATE* is significantly different from zero at the 10% level in the female wage equation but only at the 25% level in the male wage equation. Other studies have not considered the effect of unanticipated unemployment on the female wage. Heckman and MaCurdy found a negative effect of the current local unemployment rate on the female wage. However, Adams found a significant positive effect of unanticipated unemployment rate on the male wage. Thus, our results differ slightly from other studies for the effects of unanticipated unemployment on the female

real wage.⁶ This result is reasonable if married males on average have significantly larger investment in firm-specific human capital than married females. Workers having firm-specific human capital are less likely to be laid off (Becker).

Wage rates of males and females differ because of local differences in cost of living and locational amenities. Under the null hypothesis that the six coefficients of these variables are jointly equal to zero, the sample value of the *F*-statistic is 19.8 for males and 8.9 for females. The critical value of the *F*-statistic with 6 and infinite degrees of freedom at the 1% significance level is 2.1. Thus, the null hypothesis of no effect is rejected.

Cost of living differences matter. A larger land price increases significantly the wage rates of males and females. The elasticities are .060–.073 for males and .053–.056 for females. These magnitudes fall in the range predicted by Kenny and Denslow for wage adjustments to compensate for home-site cost differences when housing costs are 15% of consumption expenditures and the price of the home site accounts for 21% of the value of a home, including site. An increase of the urban share of the population increases significantly the wage rate of males. However, *URBAN* does not statistically significantly affect female wages.

Climate is a proxy for locational amenities, and the effects of *JAN* and *JULY* on wage rates are conditioned by whether the regional dummy variables are included. Because *JAN* and *JULY* are correlated with the regional dummy variables, the best measure of winter and summer temperatures on wage rates is obtained when the regional dummy variables are excluded. The effect of *JAN* on \log_e male wage is quadratic and statistically significant. In the female wage equation, *JAN* and *JAN*² do not have statistically significant effects. The effects of *JULY* on \log_e wage are quadratic and statistically significant for males and females. For males, an increase of *JULY* first causes a reduction of the wage but at a decreasing rate until it reaches the low point at 82°F which is the maximum observed value. This type of relationship was also found by Hoch and Drake. For females, the quadratic effect goes in the opposite direction. The peak occurs at 72°F

which is near the bottom of the observed values and corresponds to average July temperatures in Pennsylvania and South Dakota. Thus, for most of the observed values of *JULY*, an increase of *JULY* reduces the wage rates of males and females.

Participation in Wage Work

The bivariate probit estimates of the equations explaining the probability of wage work for farm and rural nonfarm couples are reported in table 5. The equations were fitted to 32,662 observations on rural nonfarm households and 5,865 observations on farm households. Marginal effects of the regressors on the probability of wage work are evaluated at the sample mean and reported in table 6.

The first two columns of table 5 present results for rural nonfarm couples; the last two columns are for farm couples. The estimated cross-equation correlation coefficient of the disturbances in the participation equations is positive—0.26 for farm couples and 0.19 for rural nonfarm couples—and significantly different from zero at the 1% level. These results imply: (a) that the random disturbances in married male and female wage-work participation decisions are affected in the same direction by random shocks (or unmeasured effects), and (b) that the wage-work participation decisions of married males and females are not statistically independent.

For farm and nonfarm males and rural nonfarm females, the life-cycle effect on probability of wage work is quadratic. At young ages, a higher age increases the probability of wage work. The maximum effect occurs at age 26.2 and 33.2 for farm and rural nonfarm males, respectively, and at age 20.8 for nonfarm females. At older ages, the probability of wage work decreases as age increases. For farm females, the probability of wage work is largest at a young age and may be dominated by cohort effects. Labor force participation rates of married women in younger cohorts are significantly higher than in older cohorts (Killingsworth and Heckman).

A husband or wife who has more schooling has a higher probability of wage work. Additional schooling raises an individual's market wage by more than it raises his/her reservation wage. For nonfarm females these results for females are similar to those of Heckman and of Nakamura and Nakamura. For farm males, the

⁶ The null hypothesis that the 5 coefficients of the local labor market variables in the labor demand equation for females are jointly equal to zero cannot be rejected at the 1% significance level. For equation (3), table 4, the sample value of the *F*-statistic is 1.14, and the critical value for 5 and infinite degrees of freedom is 3.02 at the 1% significance level.

Table 5. Bivariate Probit Estimates of Wage Labor Participation Equation for U.S. Farm and Rural Nonfarm Married Couples, 1978-79, 1981-82

Variables	Wage Work			
	Rural Nonfarm		Farm	
	Husband	Wife	Husband	Wife
Individual/household				
AGEM	.105 (26.82)	.027 (7.52)	.033 (3.78)	.003 (0.30)
AGEM ² /100	-.158 (41.07)	-.058 (19.69)	-.063 (7.28)	-.036 (3.83)
EDM	.044 (11.70)	-.016 (5.43)	.010 (1.31)	-.010 (1.38)
EDF	.007 (1.71)	.095 (25.75)	-.029 (3.31)	.079 (9.05)
RACE	.117 (3.55)	.333 (11.20)	.350 (3.23)	.403 (3.75)
KIDS06	-.026 (1.31)	-.497 (35.55)	-.028 (0.87)	-.386 (11.41)
KIDS618	-.008 (0.83)	-.054 (10.99)	-.037 (2.23)	-.072 (4.33)
ln ASSETINC	-.338 (4.26)	-.829 (10.57)	-.347 (2.68)	-.542 (4.08)
Local labor market conditions				
PJOBGR	.041 (2.87)	.052 (4.50)	.075 (3.31)	.036 (1.60)
PURATE	-.017 (2.55)	-.024 (4.25)	.050 (2.25)	.008 (0.37)
ΔSHRSE	-.001 (0.21)	.003 (0.55)	-.020 (1.23)	-.007 (0.40)
ESHOCK	.005 (0.71)	-.004 (0.16)	-.014 (0.91)	-.017 (1.12)
RURATE	-.013 (1.07)	-.012 (1.18)	-.004 (0.13)	-.032 (1.03)
Cost of living and locational amenities				
ln PLAND	.168 (5.97)	.040 (1.72)	-.066 (1.03)	.027 (0.41)
URBAN	-.078 (0.79)	-.190 (2.47)	-.188 (0.62)	-.094 (0.31)
JAN	.007 (1.05)	-.020 (3.99)	.053 (4.42)	.022 (1.80)
JAN ² /100	-.032 (4.35)	.015 (3.08)	-.077 (3.59)	-.010 (0.45)
JULY	-.115 (0.84)	.498 (4.55)	.081 (0.45)	-.027 (0.15)
JULY ² /100	.082 (0.90)	-.321 (4.38)	-.091 (0.75)	.014 (0.12)
Agricultural prices and climate				
ln PCROP			-.251 (1.33)	-.135 (0.71)
ln PLIVE			-.488 (1.78)	-.054 (0.19)
ln FARMWAGE			.890 (3.04)	-.291 (0.98)
ln POTINP			-.520 (0.84)	-.589 (0.93)
RAIN			.014 (1.19)	.004 (0.30)
GDD/1,000			.323 (2.82)	-.014 (0.12)
RAIN × GDD/1,000			-.038 (1.61)	-.013 (0.55)

Table 5. Continued

Variables	Wage Work			
	Rural Nonfarm		Farm	
	Husband	Wife	Husband	Wife
Regional dummies and trend				
NC	.069 (1.51)	.004 (0.12)	-.102 (0.74)	-.156 (1.15)
SOUTH	-.019 (0.39)	-.115 (2.93)	.038 (0.24)	-.314 (2.00)
WEST	.007 (0.10)	-.047 (0.80)	-.287 (1.61)	-.247 (1.40)
TIME	-.008 (0.77)	.018 (2.09)	.000 (0.02)	.007 (0.28)
Intercept	6.702 (1.30)	-11.190 (2.70)	-.030 (0.00)	5.058 (0.75)
Cross-equation correlation coeff.		.187 (15.32)		.262 (12.16)
Sample size	32,662		5,866	

Table 6. Marginal Effects (Percentage Point Changes) on the Probability of Wage Work

Regressors	Rural Nonfarm		Farm	
	Husband	Wife	Husband	Wife
Individual/household				
AGEM	-.42	-.36	-.46	-.49
EDM	.41	-.16	.15	-.15
EDF	.07	.91	-.43	1.16
RACE	1.12	3.20	5.16	5.93
KIDS06	-.25	-4.77	-.41	-5.70
KIDS618	-.08	-.81	-.55	-1.06
ln ASSETINC	-3.25	-7.96	-5.11	-7.98
Labor market conditions				
PJOBGR	.40	.50	1.10	.53
PURATE	-.17	-.24	.73	.12
ΔSHRSER	-.01	.03	-.30	-.10
ESHOCK	.05	-.01	-.20	-.25
RURATE	-.13	-.11	-.06	-.47
Cost of living and locational amenities				
ln PLAND	1.61	.38	-.98	.39
URBAN	-.75	-1.82	-2.77	-1.39
JAN	-.15	-.07	.11	.24
JULY	.10	-9.46	-.81	-.09
Agricultural prices and climate				
ln PCROP			-3.70	-1.99
ln PLIVE			-7.19	-.80
ln FARMWAGE			13.11	-4.28
ln POTINP			-7.67	-8.69
RAIN			.21	.05
GDD			4.76	-.21
Regional dummies and trend				
NC	.67	.04	-1.51	-2.30
SOUTH	-.18	-1.11	.56	-4.63
WEST	.07	-.45	-4.23	-3.64
TIME	-.08	.17	.01	.10

results are similar to those of Sumner and of Huffman and Lange. The marginal effect of a year of female schooling on the probability of wage work is larger than that of male schooling (.41 versus .91 for nonfarm males and females and .15 versus .16 for farm males and females). Thus, a year of schooling seems to increase the difference between a wife's reservation and market wage relatively more than for her husband. This is consistent with results reported by Huffman and Lange for a different data set. Also, the marginal effect of schooling on the probability of wage work for a given gender is larger for rural nonfarm than for farm adults.

Although cross-person effects of education are seldom included (e.g., Heckman; Nakamura and Nakamura; Heckman and MaCurdy 1980, 1982; Sumner), negative and statistically significant effects of a farm or rural nonfarm husband's schooling on his wife's participation and of farm wife's schooling on her husband's wage work participation occur. Huffman and Lange also found a similar negative effect of a farm wife's schooling on her husband's probability of wage work. For the rural nonfarm male, additional schooling of the wife tends to increase the probability of working for a wage.

Additional children at home under age 18 have well-known and statistically significant negative effects on the probability of wage work by married females. The largest reduction occurs for additional children under age six—about 5% per child. For older children, the negative marginal effect is larger for farm than for rural nonfarm married females. The coefficients of *KIDS06* and *KIDS618* are negative in both participation equations for married males, but they are generally weaker statistically. These results imply that additional children at home raise the reservation wage of married women relatively more than the reservation wage of married men.

Although no other study has examined the effects of local labor market conditions on the probability of wage work, the effects of these variables are largely as expected. Anticipated variables have statistically stronger effects than unanticipated ones. Higher local employment growth raises the market wage of males and females, and it increases the probability of wage work. These effects are significantly different from zero at the 5% level, except for farm wives. Recall that the labor demand equations showed that a 1% rise of the expected unemployment rate caused the male wage rate to rise by 1.1% to 1.2% but the female wage rate to rise by only 0.3% to 0.4%. Thus, if households are risk neu-

tral, the positive effects of *PURATE* on the participation of farm and nonfarm males are consistent. There is no significant effect, however, on the probability of participation of farm females. A larger Δ *SHRSER* does not significantly affect the probability of wage work of males or females, although it does increase the wage rate of males. The relative employment shocks and unanticipated unemployment rate do not have statistically significant effects either.

The positive and statistically different from zero effect of *PLAND* on the probabilities of wage work implies that market wage effects are larger than reservation wage effects. The effects of January and July temperatures on the probability of wage work are mixtures of effects through the market wage and reservation wage. Statistically significant effects on participation occur for nonfarm females and farm males. For farm males, a higher *JAN* first increases the probability of wage work; the peak effect occurs at 26°F; and then it decreases.

For farm households, the farm output and input price indexes have consistent signs, and their marginal effects are larger and stronger in the husband's participation equation. A drop of crop and livestock output prices increases the probability of wage work of husbands and wives. A higher farm wage increases the probability of off-farm wage work by husbands. Other input prices have negative coefficients, but they are not significantly different from zero. Agricultural climatic effects are important for farm males. Larger *RAIN* or *GDD* has a negative impact on labor force participation when these effects are evaluated at the sample mean.

Except for women living in the South, the regional effects are not significantly different from zero. Married women living in the South, however, have a lower probability of wage work than women in the Northeast. Other things equal, the reduction is about 1% for rural nonfarm women and 5.5% for farm women. The probability of wage work for rural nonfarm women has a positive and statistically significant trend. None of the other coefficients of *TIME* is significant.

Conclusions and Implications

This study has shown that wage premiums exist for localities anticipating labor demand growth, higher unemployment rates, larger shares of employment in services, and higher costs of living. Unanticipated negative disturbances in local labor markets associated with employment growth

and unemployment reduce wage rates. Male wage rates are more responsive to anticipated changes in local labor market conditions, and female wage rates are more responsive to unanticipated changes. The differences by gender for anticipated effects reflect less time spent by females in the labor force. Women's wage rates bear a greater burden of negative unanticipated labor market outcomes because they are more geographically and occupationally immobile.

Labor force participation decisions of households are also affected by changes in anticipated local economic conditions. For farm households, the probability of wage work increases when expected farm output prices decline and decreases when local labor demand grows. Given the general trend toward larger farms, the results show that existing agriculture is significantly easier for farm households that reside in localities where nonfarm employment is expected to grow rapidly. These effects, however, pull more strongly on males than females. However, participation by farm and nonfarm couples is unaffected by unanticipated labor market outcomes.

The main policy recommendations are that (a) public information be made available to firms and households about anticipated changes in all local labor markets, (b) regionally targeted stabilization policies be employed to moderate unanticipated changes in local labor demand, and (c) tax credits or other incentives be given to workers to hasten their movement from slow growing regions and occupations to rapidly growing ones. Future research could extend this analysis through the late 1980s when national economic conditions and local labor fluctuations were moderate, and examine the effects of local economic conditions on hours of wage work through effects on wage rates.

[Received January 1989; final revision received October 1990.]

References

- Abowd, John, and O. Ashenfelter. "Anticipated Unemployment, Temporary Layoffs, and Compensating Wage Differentials." *Studies in Labor Markets*, ed. S. Rosen, Chicago: University of Chicago Press for National Bureau of Economic Research, 1981.
- Adams, James D. "Permanent Differences in Unemployment and Permanent Wage Differentials." *Quart. J. Econ.* 100(1985):29-55.
- Becker, Gary S. *Human Capital*, 2nd ed. New York: Columbia University Press for NBER, 1975.
- DaVanzo, J., D. DeTray, and D. Greenberg. "The Sensitivity of Male Labor Supply Estimates to Choices of Assumptions." *Rev. Econ. and Statist.* 58(1976):313-24.
- Economic Report of the President, 1987*. Washington DC: Executive Office of the President, 1987.
- Gerner, J. L., and C. D. Zick. "Time Allocation Decisions in Two-Parent Families." *Home Econ. Res. J.* Dec. (1983), pp. 145-58.
- Hall, Robert E. "Turnover in the Labor Force." *Brookings Pap. on Econ. Activ.* 3(1972):709-56.
- Hamermesh, Daniel S., and A. Rees. *The Economics of Work and Pay*, 3rd ed. New York: Harper and Row, 1984.
- Heckman, J. J. "Sample Selection Bias as a Specification Error." *Female Labor Supply*, ed. J. Smith, pp. 206-48. Princeton NJ: Princeton University Press, 1980.
- Heckman, James J., and T. E. MaCurdy. "A Life Cycle Model of Female Labor Supply." *Rev. Econ. Stud.* 47(1980):47-74.
- . "Corrigendum on a Life-Cycle Model of Female Labor Supply." *Rev. Econ. Stud.* 49(1982):661-66.
- Hoch, I., and J. Drake. "Wages, Climate and the Quality of Life." *J. Environ. Econ. and Manage.* 1(1974):268-95.
- Huffman, W. E., and R. E. Evenson. "The Development of U.S. Agricultural Research and Education: An Economic Perspective," part 4. Dep. Econ., Staff Pap. No. 174, Iowa State University, Dec. 1989.
- Huffman, W. E., and Mark Lange. "Off-farm Work Decisions of Husbands and Wives." *Rev. Econ. and Statist.* 71(1989):471-80.
- Israeli, O. "Differentials in Nominal Wages and Prices Between Cities." *Urban Stud.* 14(1977):275-90.
- Johnston, J. *Econometric Methods*, 3rd ed. New York: McGraw-Hill Book Co., 1984.
- Kenny, Lawrence W., and D. A. Denslow, Jr. "Compensation Differentials in Teachers' Salaries." *J. Urban Econ.* 7(1980):198-207.
- Killingsworth, M., and J. Heckman. "Female Labor Supply: A Survey." *Handbook of Labor Economics*, ed. O. Ashenfelter and Layard. New York: North-Holland Publishing Co., 1986.
- Lilien, David. "Sector Shifts and Cyclical Unemployment." *J. Polit. Econ.* 90(1982):777-93.
- Mincer, Jacob. "Family Migration Decisions." *J. Polit. Econ.* 86(1978):749-74.
- Mincer, Jacob, and H. Ofek. "Interrupted Work Careers: Depreciation and Restoration of Human Capital." *J. Human Resour.*, Winter (1982), pp. 3-24.
- Mroz, Thomas A. "The Sensitivity of An Empirical Model of Married Women's Hours of Work to Economic and Statistical Assumptions." *Econometrica* 55(1987):765-99.
- Nakamura, A., and M. Nakamura. "A Comparison of the Labor Force Behavior of Married Women in the United States and Canada with Special Attention to the Impact of Income Taxes." *Econometrica* 49(1981):451-89.
- Rosen, Sherwin. "The Theory of Equalizing Differences." *Handbook of Labor Economics*, vol. 1, ed. O. Ashenfelter and R. Layard, pp. 641-92. New York: North-Holland Publishing Co., 1986.

- Strauss, John. "The Theory and Comparative Statics of Agricultural Household Models: A General Approach." *Agricultural Household Models*, ed. I. Singh, L. Squire, and J. Strauss, pp. 71-91. Baltimore MD: Johns Hopkins University Press, 1986.
- Sumner, Daniel A. "The Off-Farm Labor Supply of Farmers." *Amer. J. Agr. Econ.* 64(1982):499-509.
- Tolley, G. S. "The Welfare Economics of City Bigness." *J. Urban Econ.* 1(1974):324-45.
- Topel, Robert H. "Local Labor Markets." *J. Polit. Econ.* 94(1986):S111-43.
- U.S. Department of Agriculture, Economic Research Service. *Rural Economic Development in the 1980's: A Summary*. Agr. and Rural Econ. Div., Agr. Info. Bull. No. 533, Washington DC, Oct. 1987.
- U.S. Department of Agriculture, U.S. Department of Commerce. "Mean Growing Degree Days Accumulated Weekly March 1 to Indicated Dates." *Weekly Weather and Crop Bull.*, April (1970), pp. 12-16.
- U.S. Department of Commerce. *1978 Census of Agriculture*, vol. 1, part 51. Washington DC, 1980.
- . *1980 Census of Population: United States Summary*, part 1, Chap. A, Washington DC, 1981.
- . *Weekly Weather and Crop Bull.*, 29 March 1971.
- . *Weekly Weather and Crop Bull.*, 7 July 1981.
- Weiss, Michael D., M. W. Whittington, and L. D. Teigen. *Weather in U.S. Agriculture*. Washington DC: U.S. Department of Agriculture, Econ. Res. Serv. Statist. Bull. No. 737, Dec. 1985.

The Impact of Wage Differentials on Choosing to Work in Agriculture

Jeffrey M. Perloff

Based on a model of industry choice and wage determination, a 1% increase in the relative wage in agriculture increases the probability that a nonurban male with no more than a ninth-grade education works in agriculture by 3.37% at the sample mean and by 1.3% when averaged over the entire sample. A 10% increase in wages may increase the proportion of those males who choose to work in agriculture by nearly a quarter.

Key words: agriculture, industry choice, labor supply, sample selection, wages.

The likelihood of nonagricultural workers joining the agricultural work force in response to an increase in the agricultural wage is estimated in this study. Knowing the responsiveness of the labor supply to wage differentials is important for evaluating many public policies. For example, if the Immigration Reform and Control Act of 1986 (IRCA) eventually restricts the supply of ineligible immigrant labor in the United States, many farmers and legislators fear large wage increases, which will lead to significant crop losses (at least in the short run) or mass noncompliance with the law. How realistic are the fears that large wage adjustments will be required to equilibrate the hired farm worker labor market is assessed in this study.

The study is based on 1988 data from the U.S. Department of Labor, Bureau of Labor Statistics Current Population Survey (CPS), which is a random sample covering all sectors of the economy. The decision to work in the agricultural or the nonagricultural sectors and wages in each sector, controlling for choice of sector, are simultaneously determined. Then the empirically estimated model is used to simulate the increase in the share of agricultural workers from a given increase in the relative agricultural wage.

In the first section, the basic modeling meth-

odology is developed. The data set is described and summary statistics for the key variables are presented in the second section. The empirical results are discussed in the third section. Simulations are used in the next section to show the likely response of workers to higher wages in the agricultural sector. In the final section, inferences and conclusions are drawn from the analyses.

Methodology

A model of industry choice and wages by industry is used to examine the effect of relative wages and nonwage factors on industry choice. The model is a variant of those of Lee, Willis and Rosen, Nakosteen and Zimmer, and Robinson and Tomes. In this model, the natural logarithms of the hourly earnings ("wages") in the agricultural (w_a) and nonagricultural (w_n) sectors are a function of demographic and individual characteristics (X_a and X_n) and unmeasured sources of individual differences (ε_a and ε_n), and the impact of individual characteristics on wages may vary across the sectors:

$$(1) \quad w_a = X'_a \beta_a + \varepsilon_a, \text{ and}$$

$$(2) \quad w_n = X'_n \beta_n + \varepsilon_n.$$

An individual's wage in a given sector is observed only if the individual is working in that sector at the time of the CPS interview.

An individual's choice of sector may depend on nonwage factors as well as the relative wage in the two sectors. Agricultural work is more physically taxing and dangerous than many other types of work. However, many people prefer

Jeffrey M. Perloff is a professor, Department of Agricultural and Resource Economics, University of California, Berkeley.

Giannini Foundation Paper No. 986.

The author is grateful to the California Employment Development Department, Employment Data & Research Division, the Giannini Foundation, and the Harris Trust for support. He also thanks Phil Hardiman and John Milat of EDD; Howard Rosenberg, who was extremely helpful at various stages of this project; Max Leavitt, who provided expert programming assistance; two anonymous referees; and Lien Tran, who was an able and helpful research assistant.

working outside in agriculture to an indoor job. Let c be the cost or benefit (disutility or utility) of working in agriculture relative to working in another industry. This (unobserved) variable is a function of a worker's characteristics, Z :

$$(3) \quad c = Z'\delta + \varepsilon_c.$$

The worker compares this cost or benefit to the relative wage in agriculture. The wage ratio (R) between agricultural and nonagricultural sectors is approximated by the difference in the natural logarithms of these wages:

$$(4) \quad R \approx w_a - w_n.$$

The difference in the logs is a close approximation of the ratio for small differences. Alternatively, one can define R as the log wage difference so that equation (4) holds exactly.

The worker chooses to work in agriculture (industry choice, i , equals 1) only if the total benefit to working in agriculture, $R - c$, is positive:

$$(5) \quad \begin{aligned} i &= 1 && \text{if } R - c > 0, \text{ and} \\ i &= 0 && \text{if } R - c \leq 0. \end{aligned}$$

If $i = 1$, then the observed wage is w_a ; otherwise, the observed wage is w_n .

The disturbance terms (ε_a , ε_n , and ε_c) in equations (1), (2), and (3) are assumed jointly normally distributed. As a result, a probit estimation technique can be used to estimate the choice of industry equation (5). Substituting for R in equation (5) using equations (4), (1), and (2) and for c using (3), we obtain a reduced-form industry choice equation,

$$(6) \quad \begin{aligned} i &= 1 \text{ if } X'_a\beta_a - X'_n\beta_n - Z'\delta - \varepsilon_c > 0, \text{ and} \\ i &= 0 \text{ if } X'_a\beta_a - X'_n\beta_n - Z'\delta - \varepsilon_c \leq 0, \end{aligned}$$

which can be estimated using probit with the exogenous variables (X_a , X_n , and Z) in equations (1), (2), and (3) on the right-hand side. Conditional on industry choice, as determined by equation (6), the hourly earnings equations (1) and (2) can be estimated using Heckman's technique to compensate for sample selection bias. If equations (1) and (2) were estimated using ordinary least squares techniques, the estimates would be biased because workers are not randomly assigned to the agricultural and nonagricultural sectors. That is, the unobserved individual difference or disturbance terms (ε_a and ε_n) would not be normally distributed if we examine data for only those workers observed in each sector.

Based on the consistently estimated wage equations, (1) and (2), the estimated wage differential is $\hat{R} = X'_a\hat{\beta}_a - X'_n\hat{\beta}_n$. That is, although we observe a worker's wage in one sector only, these equations can be used to determine the relative wage. Substituting this estimated value for R in equation (5), the structural probit for industry choice can then be estimated.

The key exogenous variables represent regions and demographic characteristics. Because agricultural and nonagricultural wages differ geographically (reflecting differences in labor demand and supply and the type of work), regional dummies are included in X_a and X_n . Years of schooling and years of experience are also hypothesized to influence wages and are included in X_a and X_n . Wages and the costs of working in agriculture may be affected by workers' racial and ethnic characteristics because of the discrimination or because they reflect language skills and legal status. They may, for similar reasons, also affect the costs of working in agriculture; hence they are included in X_a , X_n , and Z .

Also included is a dummy variable for "married, living with spouse," which has an ambiguous effect because married couples make joint employment and housing decisions. Given the relatively high variance in agricultural wage, one may be more willing to work in agriculture if one's spouse has a steady income; however, offsetting that effect, to the degree that migration is required in agricultural jobs, living with a spouse may be difficult. Similarly, having children may make migratory living relatively unattractive; however, whole families may work together in agriculture. Thus, marital status and children are included in Z , but their effects are also uncertain.

The reduced-form probit and the wage equations can be estimated without any further restrictions. The identifying restrictions that allow us to estimate the structural probit equation are that education and experience variables affect wages but do not affect the choice of sector except indirectly through their effects on the wage differential. Thus, in the structural probit, only other demographic characteristics and the wage differential are included. It is hard to justify why a year or two more of primary education would affect one's utility from working in agriculture rather than nonagriculture except indirectly through its effect on relative wages. It is similarly difficult to see why experience would affect utility directly; however, a case might be made that age (the experience variable is a lin-

ear function of age) affects utility differently in the two sectors. Not using the experience variables as identifying restrictions, however, has only a minor quantitative (and not a qualitative) effect on the results reported below.

The Data

Data used in this study are from the CPS for the 1988 calendar year, which is a random sample of individuals by housing units throughout the United States conducted monthly over the year.¹ Because selection is based on location, agricultural workers, nonagricultural workers, the unemployed, and documented and undocumented immigrants are surveyed.

The inhabitants of any given housing unit are asked questions about economic issues at two times, separated by a four-month interval. To prevent double-counting some workers (people initially interviewed in the first eight months of the year have a second interview in the same calendar year), only the first of these interviews is used. Only individuals at least sixteen years old, with no missing variables, who usually work at least fifteen hours per week (i.e., the employed), and earned at least \$2 an hour (to eliminate observations with implausible hourly earnings) were included.

Sample Restrictions

In the following analysis, the sample is restricted to only a subset of males who are relatively likely to consider agricultural employment: workers with no more than a ninth-grade education who live outside of major cities. Because less than 2% of all CPS workers are in

agriculture, the industry choice equation for the entire sample involves the tail of the sector-choice equation's error distribution; so the results are more sensitive to the choice of the error distribution (e.g., normal versus logistic) than if the mean were closer to the center of the distribution as reported below. Estimating a probit for all workers indicates that the only important determinant of agricultural employment is gender, and the estimated equation predicts (with nearly 99% accuracy) that everyone works in nonagriculture. The wage equations are not substantially different, however, if a larger sample is used.

Education was restricted because it would be a nonproductive exercise to calculate the wage differential necessary to induce a brain surgeon or other skilled worker to start working as an agricultural field hand. Although there are some hired farm workers who are highly educated (about one in 20 are college educated), the majority of agricultural workers have less than a high school education. The average number of years of education of seasonal agricultural workers is ten according to the CPS and seven according to the Department of Labor's National Agricultural Workers Survey for the same time period. The results reported below are not sensitive to the arbitrary ninth-grade cutoff. Similar qualitative results are obtained for thresholds at eighth grade, tenth grade, eleventh grade, and the twelfth grade (short of a diploma).²

Similarly, it is unlikely that workers in the middle of Manhattan are likely to switch to farm work with any plausible agricultural wage increase; so, because we are concerned with the short-run effect of a wage differential on choice of working in agriculture, the sample has been further restricted to those hired workers who live outside of major metropolitan areas with more than 100,000 people ("nonurban" areas). Although some agricultural workers with limited education live in cities, especially in California (e.g., Fresno), over two-thirds do not. Of workers with nine years or less of schooling, only one in ten works in agriculture of all workers compared to one in five of our nonurban sample.

¹ Some critics argue that the BLS undersamples migrant and illegal agricultural workers. Although this complaint may be valid, the CPS data set is used for four reasons. First, the CPS is the only random sample that includes both agricultural and nonagricultural workers in sufficient quantities to conduct a study of individual supply responses to wage differentials. Second, the included variable "Mexican" (as opposed to Mexican American, Chicano, or other Hispanic or non-Hispanic workers) is a proxy for legal status. This proxy is imperfect, however, because some of these Mexican workers may have legal status and others may misreport whether they are Mexicans. Third, if IRCA or other programs prevent undocumented aliens from working in agriculture, one would be primarily interested in the determinants of industry choices of workers with legal status. Fourth, the empirical study reported below was also estimated using only non-Hispanic workers, and the qualitative results for all equations and the quantitative results for the wage equations are virtually the same as those reported below. Thus, it seems unlikely that including more workers without legal status in the sample or including a variable for legal status would greatly change the results.

² It is unlikely that restricting our sample to relatively uneducated workers creates a sample selection bias for two reasons. First, in the short run, education must be viewed as a predetermined variable, and selecting on a predetermined variable does not cause a sample selection bias. Second, a Heckman sample selection test (based on a sample of all workers) does not indicate a sample-selection bias.

Females were dropped because there are not enough women in the sample to estimate an agricultural wage equation for females. There are only 396 females with 12 in agriculture (and more than 12 variables in the wage equations) in the sample. The effects of also including city dwellers and females are briefly discussed below.

Means and Standard Deviations

Presented in table 1 are the means and standard deviations (for continuous variables) of several key variables for our sample of 931 men, of whom 19.4% identified themselves as hired agricultural workers (including managers and fore-

men). The table also shows that more than a quarter (26.2%) of those in the sample live in the South Atlantic region, whereas only slightly more than a fifth (21.0%) of the agricultural workers live in that region. Similarly, where only 7.7% of the sampled live in the West (Pacific region), over a quarter (27.5%) of the agricultural workers live there. Moreover, only 5.7% of those in the total sample, but 27.1% of the agricultural workers, live in California.

Relatively more blacks and other nonwhites are in the agricultural subsample than in the overall sample. Agricultural workers are less likely to be married and living with their spouses than nonagricultural workers (52.5% vs. 77.9%), perhaps reflecting the migratory nature of many

Table 1. Means and Standard Deviations

Variable	Number of Observations	All 931	Agriculture 181	Non- Agriculture 750
Binary (0-1) Variables (%)				
<u>Region</u>				
New England (CT, MA, ME, NH, RI, VT)		8.8	2.8	10.3
Mid-Atlantic (NJ, NY, PA)		4.4	0.6	5.3
East North Central (IL, IN, MI, OH, WI)		7.7	3.3	8.8
West North Central (IA, KS, MN, MO, ND, NE, SD)		10.2	14.4	9.2
South Atlantic (DC, DE, FL, GA, MD, NC, SC, VA, WV)		26.2	21.0	27.5
East South Central (AL, KY, MI, TN)		13.2	7.2	14.7
West South Central (AR, LA, OK, TX)		12.6	12.7	12.5
Mountain (AZ, CO, ID, MT, NM, NV, UT, WY)		9.2	10.5	8.9
Pacific (AK, CA, HI, OR, WA)		7.7	27.5	2.8
<u>States</u>				
California		5.7	27.1	0.5
Texas		6.2	8.3	5.7
Florida		3.7	3.9	3.6
<u>Demographic characteristics</u>				
Black		13.1	14.4	12.8
Other nonwhites		2.1	1.7	2.3
Mexican (ethnicity)		13.2	43.1	6.0
Mexican-American (ethnicity)		4.9	7.7	4.3
Other Hispanics (ethnicity)		1.4	2.2	1.2
Married, living together		72.9	52.5	77.9
<u>Job characteristics</u>				
Union member		14.7	1.7	17.9
Paid by the hour		75.8	64.6	78.5
Agricultural manager or foreman		1.1	5.5	0.0
Agriculture (hired farm workers)		19.4	100.0	0.0
Continuous Variables [mean (s.d.)]				
Number of children		.7 (1.2)	1.07 (1.5)	.6 (1.1)
Years of schooling		7.5 (2.1)	6.6 (2.2)	7.7 (2.0)
Years of experience		34.3 (13.9)	29.3 (16.9)	35.6 (12.8)
Earnings per hour ("wage")		7.13 (3.7)	4.80 (2.2)	7.69 (3.7)
Usual weekly hours		40.5 (8.6)	42.3 (10.8)	40.1 (7.9)

agricultural jobs and the relative youth of agricultural workers.

Nonagricultural workers are, on average, seven years older than agricultural workers. Because the average level of education in the United States has risen over time, non-Hispanics with relatively little formal education tend to be older than the rest of the population. The average age of workers in agriculture is forty-two; the average age of non-Hispanics is forty-nine (and the median is 50), however, the average age of Hispanics is only thirty-five (and the median is 32).

Although 17.9% of the nonagricultural workers are union members, only 1.7% of the agricultural workers are union members (all of whom are in California). In this sample, agricultural workers average one fewer year of school than nonagricultural workers. Agricultural workers also have six years less work experience (calculated as age minus 6 minus years of formal school).

Many agricultural workers receive piece-rate payments (35% receive only piece rate compared to 11% of nonagricultural workers). In this study, hourly earnings are calculated by dividing the reported weekly earnings by the reported usual weekly hours. For workers who receive time rate pay, this calculated hourly earnings number is usually identical or very close to the reported wage. The average hourly earnings of agricultural workers (\$4.80) is only 62% of those of nonagricultural workers (\$7.69). They work, on average, 2.2 more hours a week, however, so that the average weekly earnings of agricultural workers are 66% of that of nonagricultural workers (\$203 versus \$308).

The Empirical Results

The first step of the analysis is to estimate a reduced-form probit equation describing how choice of industry depends on exogenous geographic and demographic variables. Then, conditional on industry choice, wage (hourly earnings) equations are estimated for each sector. Finally, a structural probit equation is estimated.

Reduced-Form Probit Equation

The reduced-form probit (working in agriculture = 1 and working in nonagriculture = 0) coefficient and asymptotic standard error estimates are shown in the second and third columns of table 2. Because it is a reduced-form equation,

the coefficients reflect both wage differential and cost factors as described above.

Compared to the West, living in the West North Central region makes one more likely to be in agriculture (all else the same). Similarly, Californians in the sample are much more likely to work in agriculture than others in the West.

An extra year of experience makes a worker more likely to work in agriculture if the individual has at least thirty-two years of experience (and less likely otherwise). In contrast, one more year of formal schooling makes one less likely to work in agriculture if one has had at least five years of schooling. Married men living with their spouses are less likely to be agricultural workers; however, the more children one has, the more likely one works in agriculture.

Workers who report their ethnicity as Mexican (as opposed to Mexican-American, Chicano, or other non-Mexican Hispanic or non-Hispanics) are much more likely to work in agriculture than other groups. Workers who report that they are non-Mexican Hispanics (including Mexican-Americans) are also more likely to work in agriculture than other groups but not as likely as the Mexicans.

The equation correctly predicts 87.3% of the observations. The various standard probit R^2 measures (Maddala, Hensher and Johnson, Chow) range from 0.29 to 0.46.³ A log-likelihood test strongly rejects the hypothesis that only the constant term matters.

Wage Equations

The wage equations (the regression of the natural logarithm of hourly earnings on demographic and geographic variables) in table 3 were estimated using Heckman's two-stage technique to control for nonrandom industry choice. The correlation between the disturbance in the regression and the selection criteria is very high (nearly one), and the estimates of the correlation are virtually the same for the two equations. On the basis of a Heckman t -test on the selectivity parameter, we can reject the hypothesis that ordinary least squares estimates have no sample selectivity bias in both the wage equations.

The equations show a different relationship in the two sectors with respect to geographic and

³ The probit equations and summary statistics were estimated using Kenneth J. White's Shazam program, version 6.1. The sample-selection adjusted wage equations reported below were estimated using William H. Greene's Limdep program, version 5.1.

Table 2. Probit Equation: Probability of Working in Agriculture

	Reduced-Form Equation		Structural Equation	
	Coefficient	Asymptotic S.E.	Coefficient	Asymptotic S.E.
Constant	-1.6850	0.7916	-0.8320	0.6336
New England	0.8507	0.6578	1.7278	0.6718
Mid-Atlantic	0.2749	0.7558	1.8039	0.8124
East North Central	0.7512	0.6572	1.9637	0.7020
West North Central	1.6973	0.6319	2.1878	0.6269
South Atlantic	0.7273	0.6270	2.2353	0.6831
East South Central	0.8943	0.6425	2.3763	0.7039
West South Central	0.7172	0.6432	2.2399	0.7095
Mountain	0.5270	0.6298	1.9857	0.6958
California	2.5634	0.6735	4.2979	0.7616
Texas	-0.3190	0.3340	0.1003	0.3459
Florida	0.1140	0.3002	-0.1936	0.2986
Mexican	1.3387	0.2331	0.4290	0.3167
Non-Mexican Hispanic	0.9263	0.2601	0.0284	0.3384
Black	0.5120	0.1741	-0.1448	0.2335
Other nonwhite	0.3833	0.4249	-0.7369	0.4859
Married, living with spouse	-0.6666	0.1385	-0.6085	0.1298
Number of children	0.0899	0.0537	0.0825	0.0495
Years of school	0.3610	0.1403		
Years of school squared	-0.0358	0.0129		
Experience	-0.0561	0.0160		
Experience squared	0.0009	0.0002		
$R = \ln(w_a) - \ln(w_n)$	1.7273		0.3742	
Number of observations	931		931	
Log-likelihood (constant only)	-458.57		-458.57	
Log-likelihood	-300.27		-301.65	
Likelihood ratio test	316.61 (21 d.f.)		313.85 (18 d.f.)	
R^2 measures:				
Maddala R^2	0.29		0.29	
Cragg-Uhler R^2	0.46		0.46	
McFadden R^2	0.35		0.34	
Chow R^2	0.38		0.37	
Percentage of correct predictions	87.3		87.5	
	Prediction success table		Prediction success table	
	actual		actual	
	0 1		0 1	
Predicted 0	731	99	733	99
Predicted 1	19	82	17	82

demographic variables. In the nonagricultural sector, controlling for demographic characteristics, wages do not differ statistically significantly across regions or states with the sole exception that wages are higher in California. In contrast, regional differences in agricultural wages are pronounced. Compared to the West, wages are significantly lower in the New England, mid-Atlantic, East North Central, South Atlantic, East and West South Central, and Mountain regions. Surprisingly, in this sample, agricultural wages are lower in California than in the rest of the West, controlling for other factors. The wages in Texas are significantly lower than in the West.

At least for this relatively uneducated group, extra education does not statistically signifi-

cantly increase one's nonagricultural wage. In agriculture, however, extra education is positively related to wages up to five years of school but negatively for more years.⁴ Agricultural wage differences resulting from extra years of school are small (only a few pennies per hour), however.

At least until one has thirty-three years of experience, extra experience increases the non-agricultural wage; whereas, extra experience (or age) does not have a statistically significant effect on the agricultural wage. In the nonagricultural sector, having twenty years of experi-

⁴ Most previous studies based on the CPS (without educational limits on the sample) find that education does not have a statistically significant effect in the agricultural sector (e.g., Perloff) but does in the nonagricultural sector.

Table 3. Logarithmic Wage Equations Adjusting for Sample Selectivity

	Agriculture		Non-Agriculture	
	Coefficient	Asymptotic S.E.	Coefficient	Asymptotic S.E.
Constant	1.3161	0.3586	1.4641	0.1777
New England	-0.5338	0.2736	0.0637	0.1235
Mid-Atlantic	-0.9429	0.3073	-0.0006	0.1322
East North Central	-0.7060	0.2761	0.0578	0.1245
West North Central	-0.3139	0.2846	0.0540	0.1326
South Atlantic	-0.9556	0.2555	-0.0343	0.1184
East South Central	-1.0005	0.2658	-0.0745	0.1211
West South Central	-0.9481	0.2734	-0.0169	0.1248
Mountain	-0.8560	0.2648	0.0473	0.1211
California	-0.5710	0.2771	0.4837	0.2378
Texas	-0.3038	0.1215	-0.0474	0.0932
Florida	0.1633	0.1099	0.0314	0.0807
Years of school	0.1535	0.0555	0.0339	0.0395
Years of school squared	-0.0149	0.0050	-0.0023	0.0034
Experience	-0.0061	0.0069	0.0325	0.0057
Experience squared	0.0001	0.0001	-0.0005	0.0001
Mexican	0.6067	0.1136	0.0820	0.1004
Non-Mexican Hispanic	0.5082	0.1129	-0.0225	0.0840
Black	0.2511	0.0762	-0.1543	0.0521
Other nonwhite	0.7284	0.1866	0.0755	0.1061
Selectivity parameter	0.3867	0.0936	0.4607	0.1348
Number of observations	181		750	
Mean of ln(wage)	1.50		1.94	
Standard deviation of ln(wage)	0.35		0.44	
Standard error of the regression	0.27		0.40	
Standard error corrected for selection	0.40		0.47	
Sum of squared residuals	13.62		119.85	
R^2	0.39		0.17	
$F(20, 729)$	5.15		7.71	
Log-likelihood	-22.70		-376.50	
Log-likelihood (constant only)	-67.70		-448.41	
$\chi^2(20)$	90.00		143.82	
Squared correlation of disturbance in regression and selection criterion	0.937		0.979	

ence instead of ten is worth 18¢ more per hour.

In the nonagricultural sector, controlling for other factors, blacks earn 14% less per hour than whites; whereas there is no statistically significant wage differential between whites and other nonwhites or either group of Hispanics. In the agricultural sector, workers who report their ethnicity as Mexican earn 82% more an hour than whites; other Hispanics earn 65% more; and blacks earn 28% more.⁵ These cross-sector racial wage differentials may reflect either discrimination or other effects not otherwise captured by the exogenous variables in this equation.

For example, nonwhites and Hispanics may be more likely to work in jobs that pay premia, such as dangerous jobs or certain migrant jobs.

The qualitative results from these wage equations are similar to those of earlier studies (e.g., Perloff) that are based on CPS data but did not adjust for the sample selection effect of industry choice. The quantitative effects differ substantially, however. For example, in the wage equations shown above, the estimated ratio of wages in agriculture to those in nonagriculture averaged 31% over the sample. Based on estimates of the same equations using ordinary least squares, the comparable average wage ratio is 66%, or more than double, reflecting the failure to control for sample selection. Those people for whom the wage ratio is relatively high are more likely to work in agriculture as discussed next.

⁵ The unconditional mean hourly earnings by demographic group in agriculture are: all \$4.87 (with a standard error of 2.16), white 5.00 (2.14), black \$4.14 (2.13), Hispanic \$4.97 (1.70), and Mexican \$5.07 (1.68).

Structural Probit Equation

The structural probit equation is reported in the last two columns of table 2. To estimate it, the wage ratio, R , is approximated by the estimated difference in the natural logarithm of the wage one would earn in agriculture and in the non-agricultural sector.

In the structural probit equation, controlling for demographic and wage ratios, workers are more likely to work in agriculture in most other regions of the country than in the West. They are also more likely to work in agriculture in California than in the rest of the West. Of course, given that wage ratios vary geographically, much of the geographic difference in choice of sector is captured by the wage-ratio term.

The structural probit equation does not show a significant difference between choice of sector among whites and other racial or ethnic groups after controlling for wage ratios. That is, the preference of working in agriculture of these groups shown in the reduced-form equation is presumably captured in the structural equation by the wage-ratio term, which reflects relatively high agricultural wages for these groups.

The wage ratio has a large, statistically significant effect. A 1% increase in the relative wage in agriculture increases the probability that one works in agriculture by 3.37% at the sample mean. On average over the entire sample (Hensher and Johnson), a 1% increase in the relative agricultural wage increases the probability of working in agriculture by 1.3%.

The ratio of the estimated agricultural wage to the nongricultural wage is 0.37 for those workers in agriculture and only 0.30 for those who are not in agriculture. That is, the estimated wage ratio of those who choose to work in agriculture is nearly a quarter more than of those who choose not to work in agriculture. Thus, choosing to work in agriculture appears, in large part, to be based on a comparison of wages between the two sectors.

The structural probit has virtually the same explanatory power as the reduced-form equation. The probit R^2 measures range from 0.29 to 0.46; and the prediction success table shows that the correct sector is predicted for 87.5% of the sample.

Sensitivity Experiments

Other experiments were used to test the sensitivity of these results to the specifications used.

In none of these experiments was the key result (the effect of the wage ratio in the structural probit) substantially affected.

As a sensitivity test on the assumed error structure, the system was estimated using logit rather than probit equations (that is, the disturbances were modeled as logistic rather than normal). Although the key result was virtually the same as with the probit system, the correlation between the reduced-form logit equation and the selectivity equations' disturbance terms were estimated to be greater than one, which is, of course, impossible. For that reason, only the probit equations are reported here.

The system of equations reported above does not have a union variable on the right-hand side of any equation. Expanding this system of equations to include a union equation (with an error term that was correlated with the industry choice equation error term) and including a union dummy as an endogenous right-hand-side variable in the wage equations proved impossible to estimate (because of the very small number of union workers in the agricultural sector within this sample). Using the specification above but including union as an exogenous variable leaves the race and ethnicity coefficients in the wage equation virtually unaffected. As a result, union status was left to the residual term in the equations reported here. Alternatively this system may be viewed as quasi-reduced-form equations where the included demographic and regional (all the union members in the sample are in California) variables also explain union status.

As mentioned above, the results are not sensitive to the educational threshold. In another experiment, seasonal dummy variables (11 monthly dummies), which were included in all the equations, had coefficients that were not statistically significantly different from zero either individually or collectively in any equation. In yet another experiment, military veteran status was included in the probits. Although its coefficient was significant in the reduced-form probit, it is not included in the equations reported here because of its ambiguous interpretation. On the one hand, virtually all veterans have legal status (although the inverse does not hold); on the other hand, it may be a proxy for other factors such as intelligence or skills. Next, the dummy variable for non-Mexican Hispanic was divided into Mexican-American (or Chicano) and other Hispanics; however, these latter two variables had virtually identical coefficients.

Finally, a model was estimated using a larger sample that included city dwellers and females

but was still restricted to those with no more than nine years of education. Agricultural workers are only 6.6% of this larger sample. The same system was used except that a dummy variable for city and another for female were included on the right-hand side of all equations. Both these dummies had large, statistically significant effects in all equations; however, the other demographic coefficients were relatively unaffected. In the structural probit equation, the relative wage term remained large and had a *t*-statistic of 4.7. The predicted 4.4% shift into agriculture from a 1% increase in the relative wage at the sample mean is larger than in the model above. The original model is stressed in this paper because the estimates in the larger model are based on estimates in the tails of the normal (probit) distribution, which make inferences outside the sample, which we now examine, more speculative.

The Response to Higher Agricultural Wages

The system of equations can be used to simulate the effect of higher wages on the agricultural supply of nonurban, relatively uneducated workers. As shown in table 4, the estimated agricultural wage is only 29% of the estimated nonagricultural wage when averaged across the sample. The simulations reflect the effects of an across-the-board increase in the agricultural wage holding the nonagricultural wage constant. That is, the simulations increase the constant term in the regression on the logarithm of the agricultural wage, which is equivalent to a constant percentage increase in the agricultural wage for all workers.

In table 4, two methods are used to calculate

Table 4. Effect of an Increase in the Agricultural Wage

Increase in the Agricultural Wage (%)	w_a/w_n (%)	Percent Agricultural Workers	
		50% Rule	Average
0.00	29.38	10.63	19.45
2.00	30.04	10.96	20.07
4.00	30.63	11.82	20.69
6.00	31.21	12.35	21.31
8.00	31.80	12.78	21.94
10.00	32.39	13.10	22.57
20.00	35.34	15.47	25.76
30.00	38.28	18.69	28.96
40.00	41.23	21.91	32.15
50.00	44.17	25.35	35.30

the effect of the wage increase on the share of workers in agriculture as both are commonly used in probit studies. In the first method, a worker is assigned to the agricultural sector if the probability he works in agriculture (according to the structural probit equation) is at least 50%. Using this 50% rule, the model predicts that 11% of the workers will work in agriculture. In the second method, the probability of working in agriculture that the model predicts for each individual is averaged across all individuals (using equal weights). This average is 19%, which is (by the nature of the probit estimation technique) virtually the same as the actual percentage in the sample.

If the agricultural wage increased by 2%, the wage ratio would increase by 2.2%. Using the 50% rule, the share of workers in agriculture would increase by 3.1% (or 0.33 percentage points from 10.63% to 10.96%; whereas, using the second method, the share would increase by 3.2% (or 0.62 percentage points from 19.5% to 20.07%). If the agricultural wage were raised by 10% (with no response in the nonagricultural wage), the first method indicates a 23% increase in the share of agricultural workers; the second method predicts a 16% increase.

Simulations of larger wage increases should, of course, be viewed with substantial caution. If the point estimates are assumed to hold with larger increases, a 50% increase in the agricultural wage leads to a 139% increase in the share of agricultural workers using the first method (to a quarter of this relatively uneducated labor force) and an 82% increase using the second method.

In interpreting these simulations, recall that they reflect a national average for nonurban, relatively uneducated workers. There may also be (a presumably smaller) response by better educated or urban workers. Thus, the simulations reported here may be lower bounds on the true (larger) response. However, if IRCA were strictly enforced, driving undocumented workers out of agriculture, a substantial number of U.S. citizens and immigrants with legal documentation would have to be hired to replace them. One survey of California employers (Rosenberg and Perloff) indicates that in 1987, after the passage of IRCA, one-third of new hires were illegal aliens.

Conclusions

This paper presents a model of industry choice and wage determination. The chief result of this

analysis is that inducing more workers to switch to agriculture may not require large wage increases. Indeed, a 10% increase in wages may increase the share in agriculture of nonurban male workers with no more than a ninth-grade education by nearly a quarter. Nonetheless, in some states and in certain crops, half or two-thirds of the agricultural work force has been undocumented aliens so that larger wage increases may be required. Because this study has focused on only supply-side effects, a full analysis of the wage effects of a government policy that prevented undocumented workers from working requires a comparable demand-side analysis.

Further work on agricultural labor supply remains to be done as well. For example, this report has focused on the role of higher wages in attracting agricultural labor. In general, however, better working conditions and other benefits (such as health insurance and housing) could also attract extra workers, holding wages constant.

[Received November 1989; final revision received September 1990.]

References

- Chow, Gregory C. *Econometrics*. New York: McGraw-Hill Book Co., 1983.
- Heckman, James J. "Sample Bias as a Specification Error." *Econometrica* 47(1979):153-61.
- Hensher, David A., and Lester W. Johnson. *Applied Discrete Choice Modeling*. New York: John Wiley & Sons (a Halsted Press book); 1981.
- Lee, Lung-Fei. "Unionism and Wage Rates: A Simultaneous Equations Model with Qualitative and Limited Dependent Variables." *Int. Econ. Rev.* 19(1978):415-33.
- Maddala, G. S. *Limited Dependent and Qualitative Variables in Econometrics*. New York: Cambridge University Press, 1983.
- Nakosteen, Robert A., and Michael A. Zimmer. "Migration and Income: The Question of Self-Selection." *S. Econ. J.* 46(1980):840-51.
- Perloff, Jeffrey M. "Union and Demographic Wage, Hours, and Earnings Differentials among Californian and Other U.S. Agricultural Workers." *California Farm Labor Relations and Law*, ed. Walter Fogel. Los Angeles: Institute of Industrial Relations Monograph and Res. Ser. 41, University of California at Los Angeles, 1985.
- Robinson, Chris, and Nigel Tomes. "Self-Selection and Interprovincial Migration in Canada." *Can. J. Econ.* 15(1982):474-502.
- Rosenberg, Howard R., and Jeffrey M. Perloff. "Initial Effects of the New Immigration Law on California Agriculture." *CA Agr.* 42(1988):28-32.
- Willis, Robert J., and Sherwin Rosen. "Education and Self-Selection." *J. Polit. Econ.* 87(1979):S7-S36.

Labor Contracting and a Theory of Contract Choice in California Agriculture

Ann Vandeman, Elisabeth Sadoulet, and Alain de Janvry

We present a model of labor contracts, where seasonality, the sensitivity of output to labor quality and work intensity, and the relative advantages of labor contractors in recruitment and growers in supervision of seasonal farm workers determine the choice of employment contract. Differences in the optimal means of extracting work result in lower wages paid under labor contracting than direct hiring. We derive and estimate the probability of labor contracting and wages as functions of worker and job characteristics using data on California farm workers and employers. From estimated expected wages, our results indicate that successful unionization or reducing the flow of undocumented workers into California agriculture both would reduce contracting and increase wages.

Key words: agricultural labor, contract theory, labor contracting, sample selection bias.

Labor contracting is an old institution in California agriculture. Contractors have supplied crews to plant, hoe, thin, and harvest labor-intensive fruit and vegetable crops in the state for more than 100 years. It is also a persistent institution. Although advocates of farm labor unionization predicted that contracting would disappear with the introduction of collective bargaining, contractors have continued to increase their share of the seasonal labor market. Currently, about one-third of the production jobs on California farms are performed by workers supplied by contractors. How contractors contribute to the functioning of the labor market and to management of the labor process are the subjects of this paper.

Labor contractors are independent intermediaries who, for a fee, recruit, hire, and supervise seasonal farm workers. Their work force is composed primarily of immigrant workers with limited knowledge of their rights or ability to exercise them. Workers are often dependent on

the contractor for transportation, housing, and access to health care in addition to a paycheck. High levels of unemployment among these workers, resulting from labor demand seasonality and from the constant replenishment of the labor pool through immigration, foster this dependence. Abuse of this relationship has resulted in a rather unsavory reputation for contractors (Vaupel and Martin). Accounts of promised jobs that do not materialize; wages not paid; and overcharging for transportation, food, and housing are common. Growers also fall victim to abusive contractors who pad the payroll with phantom workers or exaggerate production figures. Reports of fraud and abuse led California to require licensing and bonding of contractors in 1939. Federal law followed suit in 1963.

What accounts for the persistence of this antiquated and apparently primitive institution in agriculture? Contractors were used originally to bring Chinese immigrant workers into agriculture, where lack of a common language was a barrier to direct employment and supervision (Fisher). Today, most foremen and field supervisors, and many growers themselves, are bilingual. Thus, while the work force still is composed primarily of immigrant workers, language is no longer a barrier between employer and employee.

Economic analyses of the labor process have sought to explain patterns in the organization of agricultural production by applying the theory

Ann M. Vandeman is an agricultural economist, Economic Research Service, U.S. Department of Agriculture; Elisabeth Sadoulet and Alain de Janvry are, respectively, a lecturer and a professor, Department of Agricultural and Resource Economics, University of California, Berkeley.

Giannini Foundation Paper No. 987.

The authors wish to thank two anonymous reviewers for helpful comments on an earlier version of this paper. The paper is an outgrowth of the first author's Ph.D. thesis, which was supported by the Giannini Foundation, the Senator Drobish Memorial Endowment Fund, and the Department of Agricultural and Resource Economics, University of California, Berkeley.

of transactions costs to particular problems of tenancy and labor contract choice, such as the payment of piece versus time rates (Stiglitz, Roumasset and Uy), permanent versus casual labor contracts (Bardhan, Eswaran and Kotwal 1985b), and alternative forms of tenancy (Eswaran and Kotwal 1985a). We employ this theory to examine the unique role of the contracting system in labor-intensive agriculture.

Previous studies of seasonal farm labor have identified labor demand seasonality as the source of low returns to farm work and problems of unstable labor supply (e.g., Holt, p. 11). We argue that the contracting system is advantageous to growers precisely because it preserves the casual nature of seasonal farm work. Contractors possess an advantage in recruiting workers for seasonal jobs because they can spread their recruitment costs over a number of short-term contracts. They are also more efficient than growers in the recruitment of new entrants to farm work, workers with little knowledge of the farm labor market or labor standards, and workers whose alternatives inside and outside of farm work are limited by their illegal status. We will show that hiring workers with these characteristics enables contractors to achieve higher levels of work intensity in certain types of jobs while paying lower wages than growers who hire labor services directly.

Contracting is, however, not appropriate to meet all of agriculture's labor needs. In particular, growers have an advantage in direct hiring for jobs which are less seasonal and which require greater care and knowledge of production.

In the next section we develop a formal model of the grower's contract decision and wage determination, where the contract alternatives are direct hiring and contracting for labor. The model is used to infer a distribution of contracts across types of jobs. Then we reverse the use of the model to derive the conditional distribution of workers between contract types in a given job and the predicted wages expressed as a function of contract type. Subsequent sections present empirical results for these two equations and use these results to calculate expected wages under alternative policy scenarios.

Labor Contracts and Labor Extraction

The functions of the contracting system are considered at two levels: in the labor market and within the labor process. Labor market functions include the recruitment and hiring of

workers. In the labor process, the objective is to induce workers to work at some optimal level of effort and quality.

The grower has two basic alternative means of access to labor: direct hiring and management of workers and purchasing labor from a labor contractor. Growers are assumed to make decisions on these alternatives at each stage of the production process of each crop, and each decision is assumed independent. Thus, they may choose to employ labor directly for one task and to contract for labor at a later stage of production. The choice of contract will depend on the activity to be performed and on the relative efficiencies of growers and labor contractors in the recruitment of workers and the extraction of work in that particular activity.

Supervision and the Cost of Job Loss

Growers and contractors have two means of extracting work from workers. One is by direct supervision, which primarily influences the quality of the work, and the other is to apply pressure to increase the quantity of work by the threat of job loss.

Supervision involves directing and coordinating the labor process; making the day-to-day production decisions regarding the time and levels of input use, including labor inputs, and their methods of application. The effectiveness of supervision depends upon the supervisor's ability to manage production. The importance of supervision to an activity, whether it be pruning vines, applying fertilizer, or harvesting garlic, will vary with the sensitivity of output quality to the labor process.

Growers have a comparative advantage in supervision over contractors because of their greater knowledge of the production process. The grower's proprietary interest in maintaining the yield potential of fixed inputs such as land, trees and vines, and in protecting investments in equipment, also gives him greater incentive to supervise effectively those activities requiring greater care. Let S be the cost of supervision per unit of labor time and γ_i a parameter representing the efficiency of supervision under contract i . Because growers are relatively more efficient in supervision than contractors, $\gamma_d > \gamma_c$ (throughout, we use $i = d$ for direct hiring and $i = c$ for contracting labor).

Employers can also affect work effort by controlling the cost of job loss, defined as the difference between the actual wage and the work-

er's opportunity wage (Shapiro and Stiglitz, Bowles). The cost of job loss affects work effort (Becker and Stigler, Stiglitz). The threat of job loss is more effective, and hence effort increases, the greater is the cost of job loss that the employer can impose on the worker. Thus, high costs of job loss are implied by the common observation that farm workers work "hard and scared." Compared to direct supervision, this indirect control of the worker's performance is more useful in influencing the quantity of work rather than the quality of work. We discuss the components and manipulation of the cost of job loss in more detail at the end of this section.

Given the other inputs, the output of each activity is a function of labor time, L , modified by the level of supervision, and by the work effort E , which is influenced by the cost of job loss. Let $\alpha_2(j)$ and $\alpha_3(j)$ represent the importance of labor time and supervision to the output of activity j . For a given contract type i , the output of activity j is

$$(1) \quad Q_{ji} = E_{ji}(\gamma_i S)^{\alpha_3(j)} L^{\alpha_2(j)},$$

where E_{ji} represents the average level of work effort among the workers hired which is affected by the cost of job loss.

The cost of job loss to a worker is the difference between two components: the actual wage paid and the opportunity wage of that worker. Therefore, employers can increase the cost of job loss by recruiting workers with low opportunity cost, or by increasing their wage. However, both of these actions, recruitment and wage strategies, are not done selectively for individual workers but apply to the whole pool of workers. The employer chooses an overall strategy of recruitment that will give him access to a certain pool of workers from which he will hire, applying no further selection among the individuals. Similarly, he chooses an overall level of incentive bonus on the wage, which is then applied to the structure of individual wages. Because workers in the recruited pool are not homogenous, their individual wages, opportunity costs, and work efforts will differ. The outcome of the recruitment and wage policies will then be measured by an average intensity of work among the pool of recruited workers. We discuss the recruitment and wage policy before aggregating them in an average level of work intensity.

Because a worker's opportunity wage is determined in part by his or her individual characteristics, employers can manipulate the cost of job loss by targeting groups of workers with

particular characteristics for recruitment and hiring. Let X be a vector of characteristics describing the worker, R be the per worker cost of recruitment, and δ_i be the relative efficiency of recruitment under contract i . The level of recruitment expenditure determines the pool from which workers are hired. The pool is identified by a frequency distribution of characteristics $f(X; \delta_i R)$. It may consist of a local community of settled farm workers, a migrant stream within the United States, a cross-border migration network, or a combination of these groups.

Farm labor contractors specialize in the delivery and management of seasonal labor. This gives them two advantages. First, their more extensive contacts within farm worker communities and migration networks provide them greater access to those segments of the labor force with lower average opportunity wages, consisting of high proportions of undocumented workers, nonunion workers, and workers with little or no experience in the U.S. farm labor market. Their second advantage is the ability to reduce recruitment costs per unit of labor time by spreading these costs over several contracts. Because contractors are relatively more efficient in recruitment, $\delta_c > \delta_d$.

The second means of increasing the cost of job loss is to increase the wage. High unemployment in farm work, because of seasonality and the constant influx of new immigrant workers, precludes an assumption of competitively determined wages in this market. Thus employers exercise some control over wage rates above the federal minimum wage. Growers have both greater flexibility and face greater pressure to use wage incentives. Competition for contracts reduces the possibility for contractors to use a positive wage incentive to extract higher levels of work effort. Growers also face a greater union threat and, because unionization reduces the risk of job loss, greater pressure to increase wages in order to maintain the effort incentive (Rebitzer, Schor and Bowles). They are identified with fields that can be picketed and in some cases with a product susceptible to boycott. In contrast, contractors have not been successfully unionized, and their mobility and lack of capital make it virtually impossible to enforce collective bargaining agreements. These were, in fact, the arguments made to exclude contractors from collective bargaining legislation.

The determination of the wage and the cost of job loss are specified formally in the following way. Let w_0 be the prevailing minimum wage rate for farm work. The actual wage paid by an

employer to a worker X is $w_{og}(X, B)^{\varepsilon_i}$, where $g(X, B)^{\varepsilon_i}$ is the ratio of the actual wage to the minimum wage for workers X and contract type i . This ratio is a function of three components: X , the vector of worker characteristics; B , the average wage bonus for the labor pool; and ε_i , the parameter representing the difference in wage flexibility and pressure under the two contract types. The greater wage flexibility and the greater pressure to increase wages under direct hiring imply that $\varepsilon_d > \varepsilon_c$.

The worker's opportunity wage is expressed as a function of individual characteristics and contract type, $w_i^*(X)$. Opportunity wages will vary by contract type because of differences in the hiring preferences of growers and contractors. For instance, unionized workers have a higher opportunity wage under direct hiring as they benefit from the union's bargaining power but a lower opportunity wage under contracting as growers explicitly seek nonunionized workers when obtaining workers through a labor contractor. The cost of job loss for the individual worker X is

$$(2) \quad C_i = w_{og}(X, B)^{\varepsilon_i} - w_i^*(X).$$

Aggregating over the pool of workers, levels of R and B in a given activity j and contract type i produce the following average labor effort:

$$(3) \quad E_{ji} = \int_x [w_{og}(X, B)^{\varepsilon_i} - w_i^*(X)]^{\alpha_1(j)} f(X; \delta_i R) dX.$$

The parameter $\alpha_1(j)$ represents the impact of work effort on the output of activity j .

The concept of the cost of job loss provides a link between the contractor's roles in the labor market and the labor process. The contractor's ability to recruit and hire workers with limited employment alternatives facilitates the process of labor extraction by increasing the average cost of job loss and, therefore, work effort.

Choice of Contract

Thus, the choices of four variables determine the proceeds and the costs of labor in the considered activity: (a) the level of employment L , (b) a supervision cost per unit of labor S , (c) a recruitment cost per worker R , and (d) an average wage bonus B . The proceeds of work are measured by the production of activity j under contract i :

$$(4) \quad Q_{ji} = \int_x [w_{og}(X, B)^{\varepsilon_i} - w_i^*(X)]^{\alpha_1(j)} f(X; \delta_i R) dX \cdot (\gamma_i S)^{\alpha_3(j)} L^{\alpha_2(j)}.$$

The average wage of the pool of workers hired under contract i is

$$(5) \quad \bar{w}_i = \int_x w_{og}(X, B)^{\varepsilon_i} f(X; \delta_i R) dX.$$

The overall cost of labor, including supervision and recruitment, is

$$(6) \quad [\bar{w}_i + S + R/\alpha_4(j)]L,$$

where the quantity $R/\alpha_4(j)$ is the recruitment cost per unit of labor time and $\alpha_4(j)$ is the average length of job j .

Under direct hiring, the grower takes all the decisions and will choose levels of L , S , R , and B that maximize his profit:

$$\text{Max}_{L,S,R,B} \Pi_{dj} = p_j Q_{dj} - (\bar{w}_d + S + R/\alpha_4(j))L.$$

Under a common form of indirect contracts, these decisions are taken by the contractor who is paid a fixed fee for a given task. His objective function in deciding on levels of L , S , R , and B is to minimize the overall labor costs, including recruitment and supervision, to perform the task \bar{Q}_j specified by the grower:

$$\begin{cases} \text{Min}_{L,S,R,B} (\bar{w}_c + S + R/\alpha_4(j))L, \\ \text{s.t. } Q_{cj} = \bar{Q}_j. \end{cases}$$

Competition among contractors ensures that their fees are set to cover just their costs, equal to the sum of the labor costs, and a nominal fee for their service. This leaves the grower with a profit:

$$\Pi_{cj} = p_j \bar{Q}_j - (\bar{w}_c + S + R/\alpha_4(j))L - \overline{CF},$$

where \overline{CF} is the contractor's service fee, and L , S , R , and B are set at their optimal levels.

As the contractor's cost-minimizing problem is formally equivalent to a profit-maximizing problem, the optimal choice of L , S , R , and B under contracting can also be written as a solution of the grower's profit maximization problem:

$$\text{Max}_{L,S,R,B} \Pi_{cj} = p_j Q_{cj} - (\bar{w}_c + S + R/\alpha_4(j))L - \overline{CF},$$

with the decision taken, however, by the contractor with his comparative characteristics ε_c , w_c^* , δ_c , and γ_c .

The final decision as to which of these contracts will be used in the hiring of labor rests on the grower, who chooses the contract which yields him the highest profit. In a world of certainty, with identical growers, the same optimal contract will always be chosen for the same activity. However, in the real world of differentiated growers, the profitability of a given contract is not identical for all growers and, thus, neither is the optimal contract. In fact, we observe a distribution of contracts for each activity, not a unique dominant contract. This distribution of contracts is analyzed next.

The Distribution of Contracts by Job Types

Let $\bar{\pi}_i$ be the expected maximum profit for contract i in a given activity (the activity subscript j is omitted). Because growers differ in management ability, their profit expectations will be distributed around $\bar{\pi}_i$. Thus, any particular grower's contract choice will be based on his expected profits expressed as

$$(6) \quad \pi_i = \bar{\pi}_i + u_i,$$

where u_d and u_c are random variables. The probability that a grower chosen at random will choose contract c is

$$(7) \quad P(c) = P(u < \bar{\pi}_c - \bar{\pi}_d),$$

which is equal to the value of the cumulative distribution of u at $(\bar{\pi}_c - \bar{\pi}_d)$, or $F(\bar{\pi}_c - \bar{\pi}_d)$. Profits are calculated for each activity, so the characteristics of the job—the importance of the quality and quantity of labor to output—ultimately determine which contract is chosen.

In this empirical analysis, the data are from a

statewide survey of farm workers conducted in 1983 (Mines and Martin). The survey sample is made up of an aggregation of samples chosen from agricultural areas served by forty-two job service center offices in California. The interviewers, job service employees who regularly work with agricultural employers and farm workers in these offices, constructed quotas for sampling within their office jurisdictions based on their knowledge of the area and on guidelines developed by the survey team. Criteria were location, ethnicity, household composition (insuring a minimum number of interviews with single person households), sex, and job type (insuring a minimum number of interviews with irrigators and equipment operators). Respondents were fourteen years of age or older, household heads or spouses, and had worked at least twenty-five days in farm work during the twenty-four months prior to the survey date. Employer type was not a criterion for sample selection. Although the sample is not random, the sample selection process should not bias the results on the distribution of jobs, workers, and wages by contract type. Only observations on workers employed in nonsupervisory positions in their current job by either a labor contractor or a grower are included. Foremen, those working in off-farm jobs, and those employed by a packing house or sharecropper at the time of the interview were excluded from the sample. The latter two types of contracts cover only a small fraction of jobs and are restricted to particular crops and regions.

Table 1 illustrates the effects of job type on contract choice among growers in California. Overall, 28% of the jobs in this sample were performed by contract labor. Contracting ac-

Table 1. The Distribution of Contracts by Job Type and Average Job Length for Agricultural Jobs in California, 1983 Sample

Job	Direct Hiring		Contracting		Average Job Length Weeks
	Number	(%)	Number	(%)	
Harvest	290	65.2	155	34.8	11.5
Prune	53	84.1	10	15.9	12.7
Thin	14	82.4	3	17.6	16.8
Hoe	47	50.5	46	49.5	10.1
Irrigate	68	93.2	5	6.8	25.3
Sort	53	69.7	23	30.3	6.3
Plant	16	84.2	3	15.8	26.3
Machine operator	107	92.2	9	7.8	21.0
Sample total	648	71.8	254	28.2	13.8

Source: University of California/California State Employment Development Department (UC/EDD) Farm Worker Survey, 1983.

counts for 35% of all harvesting jobs, 50% of hoeing (weeding and thinning row crops), and 30% of sorting. In contrast, growers hire directly for over 90% of all irrigator and machine operator positions.

Two factors are at work in determining this distribution of contracts. First, where the payoff to supervision is high because of the key role of work quality, as in irrigating or operating machinery, more direct hiring is observed. Where work intensity has the greater effect on output, as in the most routine and repetitive tasks, contract hiring is more prevalent. Timing and the effect of recruitment costs on contract choice also are evident in these results. Growers more often contract out the short season jobs, such as harvesting and sorting, where their unit costs of direct recruitment and hiring are greatest. The lower unit costs of recruitment and hiring in more permanent positions are thus another reason we see growers directly hiring most irrigators and machine operators.

The Distribution of Workers between Contract Types

The fact that, for a given activity, there is both direct hiring and contract hiring means the same task could be performed by the same worker under either one of the contracts. However, with different comparative advantages in recruitment and supervision, we have seen that growers and contractors do not recruit from the same pools of workers. For that reason, some workers are more likely to be hired directly and others to work under labor contractors. The resulting distribution of workers between the two contracts is now examined. As will be shown, it depends not only on the type of job but also on individual characteristics that determine the pool from which the worker is most likely to be recruited.

For a given job j , if contract i is chosen and R and B are set at their optimum levels, then the distribution of workers hired under that contract is $f(X|i, j)$. The total distribution of workers for a given activity j is

$$(9) \quad f(X|j) = [1 - F(\bar{\pi}_c - \bar{\pi}_d)]f(X|h, j) + F(\bar{\pi}_c - \bar{\pi}_d)f(X|c, j).$$

The probability that a worker selected at random from a given activity and with characteristics X is hired under a labor contractor is

$$(10) \quad f(c|X, j) = f(X|c, j)F(\bar{\pi}_c - \bar{\pi}_d)/f(X|j).$$

For a given job j and worker k , these probabilities are dependent on the characteristics $\alpha(j)$ of the j th job and the characteristics X of the k th worker. The distributions depend on the job specifications because of their effect on the relative profitability of the two contract types and, hence, on the choice of contract. They also depend on the worker's characteristics because of the differences in the optimal levels of recruitment and in the relative efficiencies of recruitment between contracts. Finally, these distributions depend on the relative efficiency of supervision and use of the wage incentive by growers and labor contractors. Therefore, the reduced form for the probability of contract i is

$$(11) \quad f(i|X, j) = h(X, \alpha(j)),$$

where $\alpha(j) = [\alpha_1(j), \alpha_2(j), \alpha_3(j), \alpha_4(j)]$ is a vector of characteristics of activity j .

Wages are also endogenous and jointly determined along with the choice of contract as a function of job type and worker characteristics. The theoretical form was written $w_{og}(X, B)^{e_i}$, where B , set at its optimal level, depends on the $\alpha(j)$ and on the contract characteristics $\beta(i) = [\varepsilon_i, w_i^*(X), \delta_i, \gamma_i]$. Hence, a reduced-form expression for wages is

$$(12) \quad w = f[\alpha(j), X, \beta(i)].$$

Because the cost of job loss is defined as the difference between the wage paid and the opportunity wage, observed wages should increase with opportunity wages which depend on alternative sources of wage and nonwage income and on the contract type. Wages should be relatively higher in direct hiring, where recruitment costs are higher and the threat of unionization reduces the cost of job loss.

Econometric Results

We now turn to the estimation of the reduced-form equations for the probability of labor contracting and for the wage and to the empirical verification of the assertions derived from the theoretical model. Probit approximations of the distribution of contracts (11) and linear approximation of the wage (12) are as follows:

$$(13) \quad LC_h^* = a_1Z_j + b_1X_k + c_1N_r + u_h,$$

$$(14) \quad P(LC_h = 1) = \Phi(LC_h^*), \text{ and}$$

$$(15) \quad w_h = a_2Z_j + b_2X_k + c_2N_r + eLC_h + v_h,$$

where LC_h^* is a nonobservable criterion function

for selection of the worker k , in job j , and region r by a labor contractor ($h = [j, k, r]$ is an index of the observation), LC_h is a dummy variable equal to 1 if, in the observation h , hiring is done through a labor contractor and 0 if there is direct hiring, Φ the cumulative normal density function, w_h is the wage, Z_j is a vector of job specifications, X_k is a vector describing the worker's characteristics, and N_r is a vector of regional variables. The random variables u_h and v_h are assumed distributed normally with zero mean and variance $\sigma_u^2 = 1$ and σ_v^2 , respectively.

Table 2 contains a complete list of the variables used to estimate the model and their definitions, taken from the farm worker survey data referred to above. The personal characteristics variables included in the model are sex, years of farm work experience in the United States, union membership, legal status, and migrancy status. The number of years of farm work in the

United States captures the accumulated knowledge of the U.S. farm labor market, where and how to get jobs, the wage structure, and employee rights. It is not meant to measure any qualification that would influence the productivity of the worker because such qualification requirements do not differ significantly among the types of jobs considered in this study.

Ideally, the vector of job characteristics would include a duration variable and indices for the relative importance of labor quality (supervision) and quantity (work effort) in each crop and task. However, the information available from this and other similar surveys is not sufficient to construct the necessary variables. For example, the survey data on job length are right-censored at the interview date in every case. Therefore, dummies were used for the crops and tasks. The job types included in the model are harvesting jobs, pruning, hoeing, irrigating, sorting, and machine operation. Job effects are

Table 2. Variable Definitions

Dependent variables

Wage rate:	the hourly wage rate for the worker (hourly earning for piece rate jobs).
LC:	a dummy variable for contract type = 1 if the worker is employed by a labor contractor, 0 if he or she is employed directly by a grower.

Worker characteristics (X)

Sex:	a dummy variable = 1 for males, 0 for females.
U.S. farm work experience:	years of farm work experience in the United States.
Union membership:	a dummy variable for union membership = 1 if the worker has been a union member in the last 3 years, 0 if not.
Illegal status:	a dummy variable = 1 if the worker is undocumented, 0 if he or she is legally working in the United States.
Migrant:	a dummy variable for migrancy status = 1 if the worker is away from home on his/her current job, 0 if not.

Job specifications (Z)

Harvest citrus:	a dummy variable = 1 for citrus harvesting jobs, 0 otherwise.
Harvest grape:	a dummy variable = 1 for grape harvesting jobs, 0 otherwise.
Harvest vegetable:	a dummy variable = 1 for vegetable harvesting jobs, 0 otherwise.
Harvest field fruit:	a dummy variable = 1 for field fruit harvesting jobs, such as strawberries or melons, 0 otherwise.
Prune vines:	a dummy variable = 1 if the job is pruning vines, 0 otherwise.
Hoe:	a dummy variable = 1 if the job is hoeing, all crops, 0 otherwise.
Irrigate:	a dummy variable = 1 if the job is irrigating, all crops, 0 otherwise.
Sort:	a dummy variable = 1 if the job is sorting, all crops, 0 otherwise.
Machine operator:	a dummy variable = 1 for machine operator jobs, all crops, 0 otherwise.
Time rate:	a dummy variable = 1 if the worker is paid by the hour, 0 if by the piece.
Peak employment:	a dummy variable = 1 if worker's employer hires more than 50 workers at peak, 0 otherwise.

Regions (N)

Southern California:	a regional dummy variable = 1 for workers in Southern California (including Los Angeles, San Bernardino, Orange, Riverside, San Diego, and Imperial Counties), 0 otherwise.
Coastal region:	a regional dummy variable = 1 for workers in the Coastal Region (including Contra Costa, Alameda, San Mateo, Santa Cruz, Santa Clara, San Benito, Monterey, San Luis Obispo, Santa Barbara, and Ventura Counties), 0 otherwise.
Sacramento Valley:	a regional dummy variable = 1 for workers in the Sacramento Valley (including Sacramento, Yolo, Yuba, Sutter, Colusa, Butte, Glenn, and Tehama Counties), 0 otherwise.

compared to the deciduous fruit harvest. The jobs requiring greater care in their performance and, hence, greater supervision are pruning, equipment operation, and irrigating. These jobs also last longer than other farm jobs. For these reasons, growers should find it more profitable to fill these positions through direct hiring. In contrast, because work intensity is the primary factor determining output in most harvesting operations, a higher probability of contractor employment is expected in these jobs, with some variation between crops as a function of job length and quality effects on output. Peak employment is included in the model to account for variation in the levels of union threat and recruitment cost per worker by farm size.

The method of payment follows from the grower's choice of crop and is, therefore, exogenous. Piece rates are used in jobs where work intensity is critical to output; thus, we expect the probability of contracting to be positively related to piece rate work. If piece rates also serve as a screening device (Stiglitz), then piece work should yield higher hourly earnings as the more productive workers will choose these jobs. But the gains may not be equal for workers under both contracts. Piece rate earnings are dependent in part on crop yield, and there likely is a tendency to employ contract labor in the least productive piece work positions. Suppose, then, that fields can be divided into high yield and low yield. We would expect a greater increase in earnings over hourly rate jobs in the high yield fields. However, we do not observe yield differences, only that high yield is usually directly hired and low yield usually contracted. Therefore, the difference in productivity is captured by including an interactive term for the method of payment variable (LC^* time rate).

Interaction terms also are included in the wage equation to allow the effects of worker characteristics on wages to vary by contract type. The effects will not be equal where these characteristics are viewed differently by growers and labor contractors [$w_i^*(X)$ in equation (2)].

Regional effects are normalized on the Central Valley and include dummies for Southern California, the coastal region, and the Sacramento Valley. These variables represent regional differences in recruitment costs, the effect of crop mix on labor demand seasonality, and regional variation in the influence of labor unions.

Estimation results are reported in table 3. The estimated effects of U.S. farm work experience, legal status, and union membership on the prob-

ability of contract employment (column 1) are all consistent with the theoretical model. Undocumented workers are more likely to be employed by a contractor than directly by a grower. The contractor is better able to channel undocumented workers into the farm labor market and to absorb the cost of replacing workers who are picked up and deported. Under the new sanctions against hiring undocumented workers imposed by the Immigration Reform and Control Act of 1986, we may see increasing numbers of growers choosing to contract rather than risking the cost of violation.

Union membership has a significant negative effect on the probability of contractor employment. Union members are entitled to job placement services and exclusive rights to jobs covered by union contract. Then the question is, why would union members work for contractors? Although unions provide some job security, union jobs make up only a small fraction of all farm jobs, and they are subject to the seasonality of labor demand. Therefore, union workers may choose employment with a contractor over unemployment during seasonal layoffs.

Crop/task variables distinguish labor requirements by seasonality, work quality, and optimal work intensity. The negative and significant coefficients for irrigators and machine operators confirm that growers are more likely to hire directly for jobs requiring more supervision. Positive coefficients for hoeing and vegetable harvesting are also consistent with the hypothesized effects of work quality and timing on labor demand decisions. Hoeing row crops is unskilled work performed intermittently during the growing season, and vegetable harvesting is a short-season task emphasizing work intensity in most crops. Citrus harvesting is a long-season task where quality is relatively important; yet, it is more likely to be contracted. This may reflect growers' reaction to union organizing in citrus. Many growers withdrew from the commodity associations whose members were bound by union contracts won in the late 1970s and substituted much cheaper contract labor (Mines and Anzaldúa).

The Wage Equation

We have argued that the grower's choice of contract is endogenous. Because both this decision and wage rates are dependent on the same set of characteristics, any omission or measurement

Table 3. Equation Estimates for Contractor Employment and Wages

Variable Name	Contractor Employment		Wage		Derived Wage Equation Under Contractor	
	Estimated Coefficient	t-Ratio	Estimated Coefficient	t-Ratio	Estimated Coefficient	t-Ratio
Constant	-.11	-.5	6.80	11.4	3.42	4.6
Worker characteristics						
Sex	.07	.5	.62	2.6	.12	.4
Farm work experience	-.02	-3.1	-.02	-1.6	-.07	-3.9
Illegal status	.44	3.0	.02	.1	.24	.7
Union membership	-.62	-3.4	.49	1.6	-1.24	-2.5
Migrant	.18	1.2	.38	1.5	-.03	-.1
Job characteristics						
Harvest citrus	.58	1.9	2.21	4.7	2.21	4.7
Harvest grape	-.39	-1.6	-.53	-1.5	-.53	-1.5
Harvest vegetable	.51	2.4	.75	2.1	.75	2.1
Harvest field fruit	-.41	-1.4	.54	1.3	.54	1.3
Prune vines	-.53	-1.8	.01	.0	.01	.0
Hoe	.84	3.6	1.33	2.9	1.33	2.9
Irrigate	-.83	-2.8	-.01	.0	-.01	.0
Sort	.24	1.0	.66	1.9	.66	1.9
Machine operator	-.73	-2.7	.32	.8	.32	.8
Peak employment	.34	2.8	.67	3.4	.67	3.4
Time rate	-.36	-2.5	-2.57	-10.1	-1.58	-4.5
Region						
Southern California	-.45	-2.4	-.27	-.9	-.27	-.9
Coastal region	-.70	-4.3	.63	2.0	.63	2.0
Sacramento Valley	-.43	-2.3	-.60	-2.1	-.60	-2.1
Contract type						
LC			-3.38	2.9		
Interactive variables						
LC* Sex			-.51	-1.4		
LC* Farm Work Experience			-.05	-2.9		
LC* Illegal Status			.21	.5		
LC* Union Membership			-1.73	-3.6		
LC* Migrant			-.41	-1.1		
LC* Time Rate			.99	2.8		
Selection bias			1.87	2.6		

Notes: Total observations, 699. Proportion of successful prediction in probit model, 76.4%. Adjusted R-square in wage equation, 0.32.

error in these characteristics may lead to a correlation of the error terms u and v . This would, in particular, introduce a bias in the wage equation as

$$\begin{aligned} E(w_h|LC_h = 1) &= a_2Z_j + (b_2 + b'_2)X_k \\ &\quad + c_2N_r + e + \sigma_{uv} \frac{\varphi}{\Phi}, \text{ and} \\ E(w_h|LC_h = 0) &= a_2Z_j + b_2X_k + c_2N_r \\ &\quad - \sigma_{uv} \frac{\varphi}{1 - \Phi}, \end{aligned}$$

in which ϕ and Φ are the density and cumulative normal distribution evaluated for the estimated value of the criterion function LC_h^* , and σ_{uv} is the covariance of u and v . Testing and correction for this potential selection bias is done by following the Heckman method (Maddala). It consists of an OLS estimation of an extended wage equation:

$$\begin{aligned} w_h &= a_2Z_j + (b_2 + b'_2LC_h)X_k \\ &\quad + c_2N_r + eLC_h + \sigma_{uv}\lambda + \varepsilon, \end{aligned}$$

in which λ is the Inverse Mill's Ratio, equal to ϕ/Φ or $-\phi/(1 - \Phi)$, and ε is an unbiased normal error term.

The results of the wage equation appear in column 2 of table 3. Column 3 contains derived estimates for the total effects of variables on wages under contracting. The signs of the coefficients are consistent with the model, and some significant differences occur in the determination of grower and contractor wages. Thus, the wage difference by gender in direct hiring disappears under contracting. U.S. farm work experience has a small but significant negative effect only on contract wages, and the positive effect of union affiliation under direct hiring becomes significantly negative under contracting. Legal status and migrancy do not appear to affect wage rates under either contract.

As expected, larger employers pay slightly higher wages, and the piece rate differential is larger under direct hiring. We believe the latter results from direct hiring for the most productive piece work jobs and contracting others. Some self-selection into direct hire piece work may also be involved in producing this result.

Somewhat surprisingly, net wage differences are not significant in irrigating and machine operator positions, where supervision has a greater effect on output and direct hiring dominates. The prevalence of direct hiring itself may explain the higher wages observed in these jobs. On the other

hand, wages are significantly higher in citrus, vegetable harvesting, hoeing, and sorting, activities in which work effort has a greater effect on output and unions have had some impact. Thus, in a given contract, the importance of work effort does appear to increase use of the wage incentive. Wages are higher in the Coast region, where high value vegetable crops are grown under the control of a relatively few large produce companies and unionization has been relatively successful. Wages in the Sacramento Valley, by contrast, are lower than in the Central Valley.

The average wage for all workers with given qualifications and on a given job is

$$\begin{aligned} (16) \quad E(w_h) &= a_2Z_j + (b_2 + b'_2\Phi)X_k \\ &\quad + c_2N_r + e\Phi, \end{aligned}$$

which shows the two components of the effect of labor contracting on the wage. The term $b'_2\Phi$ captures the difference in opportunity cost of a given worker X_k under contractor and grower hiring. The last term measures an additional impact varying with the probability of being hired by a contractor from 0 to the maximum value of $e = -3.38$ dollars.

This last effect comes from the difference between the wages paid to workers with the same qualifications in the same job under the two contracts:

$$\begin{aligned} E(w_h|LC_h = 1) - E(w_h|LC_h = 0) \\ = b'_2X_k + e + \sigma_{uv} \frac{\varphi}{\Phi(1 - \Phi)}. \end{aligned}$$

The coefficient e is thus an overevaluation of the negative impact of the labor contracting system if σ_{uv} is positive, and an underevaluation if the correlation is negative. Furthermore, for a positive correlation as we have found, the overall effect will be negative only for sufficiently small values of the ratio $\varphi/[\Phi(1 - \Phi)]$.

This ratio $\varphi/[\Phi(1 - \Phi)]$ can be shown to have a U-shape, with a minimum at $4/\sqrt{2\pi}$ when $LC^* = 0$, and infinite value for LC^* tending to infinity. Therefore, with positive σ_{uv} , corresponding values μ and ν exist such that

$$\begin{aligned} e + \sigma_{uv} \frac{\varphi}{\Phi(1 - \Phi)} \leq 0 &\Leftrightarrow |LC^*| < \mu \\ &\Leftrightarrow |\Phi - .5| < \nu. \end{aligned}$$

The econometric results of table 3 give $e/\sigma_{uv} = 1.81$, for which $\nu = .34$. In other words, for all workers with a probability of being hired by a contractor between 16% and 84%, their wage

under a contractor is below the wage that they could expect under direct hiring, beyond the differential opportunity cost they may have under the two systems. Because this range covers the large majority of the sample, we can conclude that wages under contracting are consistently lower than under direct hiring.

This finding supports the hypothesis that contractors extract an equivalent amount of work from their employees at a lower direct labor cost than growers are able to achieve. On the one hand, they are compelled by competitive pressure to keep wages low. On the other, their ability to recruit workers with a lower opportunity wage and their greater immunity from unionization reduce the contractor's need to rely on positive wage incentives to increase the level of work effort.

Labor Policy, Contracting, and Expected Wages

Some of the factors determining the probability of contracting and the level of wages, discussed above, are subject to influence by state policies. State and federal governments, in fact, have a history of involvement in shaping employment contracts in agriculture, most notably through the administration of the Bracero Program during and after the Second World War. California is also one of the few states with legislation establishing collective bargaining rights for agricultural workers.

The implications of immigration and labor relations policies for the contract and wage structure of the farm labor market can be explored by comparing the expected wages of workers who differ in one or more of the characteristics affecting the probability of contracting and the wage for a given job and contract type. The results of the estimation are used to calculate expected wages from equation (16).

The results are presented in table 4. The worker characteristics that vary in the table are union status and legal status. The jobs examined are deciduous tree fruit harvesting, vegetable harvesting, and hoeing, with varying payment methods and locations. Among these, the highest wages are for vegetable harvesting on the coast and the lowest for hoeing in the Central Valley.

In each of the jobs examined in the table, expected wages are lowest for undocumented workers who are not unionized. It follows that

Table 4. Expected Hourly Wages for Workers with Selected Jobs and Characteristics

	Union		Nonunion	
	Legal	Illegal	Legal	Illegal
	----- (\$) -----			
Vegetable harvest, piece rate				
Central Valley	6.86	5.91	6.08	5.55
Coastal region	8.90	8.18	8.00	7.35
Deciduous harvest, piece rate				
Central Valley	7.22	6.40	6.29	5.64
Hoe, hourly rate				
Central Valley	5.43	4.65	4.87	4.45
Coastal region	7.19	6.65	6.51	6.04

Note: Calculated for a male worker with 10 years of farm work experience and working for a large employer (peak employment = 1).

the Immigration Reform and Control Act of 1986 should have the effect of increasing their wages for several reasons. First, the legal status of workers will increase their probability of working for a grower, which also will increase their expected wage. Second, assuming effective enforcement, a decline in the supply of undocumented workers will eliminate the contractors' advantage in drawing from this labor pool, thereby increasing relative per worker costs and reducing the relative efficiency of recruitment under contracting.

Unionization increases expected wages in each of the jobs examined, regardless of legal status. This is the combined result of a greater probability of direct hiring for unionized workers and the positive union wage effect among grower employees. These factors overcome the negative wage effect of union membership among contractor employees. Moreover, collective bargaining agreements usually restrict the use of contract labor even during seasonal peaks by obligating growers to hire through the union first. Thus, policies that promote farm labor unionization are expected to reduce contracting and increase wages in farm work.

Such policies are embodied in California's Agricultural Labor Relations Act (ALRA), which establishes the right to collective bargaining and provides an institutional framework for governing labor/management relations in agriculture. However, interpretation and enforcement of the ALRA are matters of political will. Without support from powerful employer organizations and labor to bring about acceptance of collective bargaining, it is unlikely that the share of union relative to contract labor will increase.

Conclusion

Three main findings support the theoretical model of labor contract choice presented in this paper. First, the effects of supervision and seasonality are confirmed by the finding that the probability of contracting is lower in jobs such as irrigating or equipment operator positions where work quality has a greater effect on output and where the length of employment is greater. Second, the higher probability of contract employment for undocumented workers, nonunion workers, and those with fewer years of farm work experience demonstrates the advantage of contractors in recruitment of workers whose opportunity costs of seasonal farm work are low. It also reveals their preference for workers who are least likely to pose a threat to management control of the labor process. Finally, the results show that, over a wide range, wages paid by contractors are lower than wages in direct hiring, net of differences in the distribution of jobs or workers between contract types.

The production process for most crops involves activities of short and long duration and requires varying levels of labor quality. Growers can, and the model predicts that they would be likely to, combine contract types over the growing season. Such a strategy would offer a further means of decreasing unity among workers, because differences in their wages, working conditions, and relationship to the grower would tend to isolate contracted from directly hired employees. Much more could be learned regarding contract choice by observing these patterns of labor use on farms.

The history of labor relations in agriculture has been marked by a sometimes intense struggle for unionization. This movement has challenged management's exclusive control of the labor process. The outcomes of the conflict are evident in the passage of the ALRA, improvements in working conditions, and the introduction of benefits and higher wages for agricultural workers. However, they are also evident in increasing mechanization, continuing political pressures from grower organizations to weaken the ALRA, and increasing use of the contracting system to reduce the threat of unionization. Neither the union movement nor the contracting system are likely to disappear from California agriculture. However, state policies that affect labor supply, such as immigration reform, and policies that determine the rights and protections afforded organized farm labor affect the relative efficiency of contracting. Hence, they

have the potential to reduce contracting's share of the farm labor market.

[Received February 1990; final revision received October 1990.]

References

- Bardhan, Pranab K. *Land, Labor and Rural Poverty: Essays in Development Economics*. New York: Columbia University Press, 1984.
- Becker, G. S., and G. Stigler. "Law Enforcement and Compensation of Enforcers." *J. Legal Stud.* 3(1974):1-18.
- Bowles, S. "The Production Process in a Competitive Economy: Walrasian, Neo-Hobbesian, and Marxian Models." *Amer. Econ. Rev.* 75(1985):16-36.
- Eswaran, Mukesh, and Ashok Kotwal. "A Theory of Contractual Structure in Agriculture." *Amer. Econ. Rev.* 75(1985a):352-67.
- . "A Theory of Two-Tier Labor Markets in Agrarian Economies." *Amer. Econ. Rev.* 75(1985b):162-77.
- Fisher, Lloyd. *The Harvest Labor Market in California*. Cambridge MA: Harvard University Press, 1953.
- Holt, James. "Introduction to the Seasonal Farm Labor Problem." *Seasonal Agricultural Labor Markets in the United States*, ed. Robert D. Emerson, pp. 3-32. Ames: Iowa State University Press, 1984.
- Maddala, G. S. *Limited-Dependent and Qualitative Variables in Econometrics*, ed. A. Deaton, D. McFadden, and H. Sonnenschein. Cambridge: Cambridge University Press, 1983.
- Mines, Richard, and Ricardo Anzaldúa. *New Migrants vs. Old Migrants: Alternative Labor Market Structures in the California Citrus Industry*. Program in U.S.-Mexican Studies Monograph No. 9, University of California, San Diego, 1982.
- Mines, Richard, and Philip Martin. *A Profile of California Farm Workers*. Giannini Information Series No. 86-2, Div. Agr. and Nat. Resour., University of California, Berkeley, July 1986.
- Rebitzer, James B. "Unemployment, Long-Term Employment Relations, and Productivity Growth." *Rev. Econ. and Statist.* 69(1987):627-35.
- Schor, Juliet B., and Samuel Bowles. "Employment Rents and the Incidence of Strikes." *Rev. Econ. and Statist.* 69(1987):584-92.
- Roumasset, James, and Marilou Uy. "Piece Rates, Time Rates, and Teams: Explaining Patterns in the Employment Relation." *J. Econ. Behavior and Organiz.* 1(1980):343-60.
- Shapiro, Carl, and Joseph E. Stiglitz. "Equilibrium Unemployment as a Worker Discipline Device." *Amer. Econ. Rev.* 74(1984):433-44.
- Stiglitz, Joseph E. "Incentives, Risk and Information: Notes Toward a Theory of Hierarchy." *Bell J. Econ.* 6(1975):552-79.
- Vaupel, Suzanne, and Philip L. Martin. "Farm Labor Contractors." *CA Agr. Mar-April* (1986), pp. 12-15.

The Welfare Effects of Targeted Export Subsidies: A General Equilibrium Approach

Mary Bohman, Colin A. Carter, and Jeffrey H. Dorfman

A three-country model of export subsidies is developed with an exporter, an importer, and a neutral country which can act on either side of the market. When the neutral country is an exporter, the country offering a targeted export subsidy always suffers a welfare loss. However, when the neutral country is an importer, the possibility of a paradoxical result—that the subsidizing country can gain and the subsidized country can lose—is shown to exist, and the conditions under which this result occurs are derived. The model fails to provide justification (on national welfare grounds) for widespread use of targeted export subsidies such as the export enhancement program.

Key words: general equilibrium, international trade, policy, welfare.

This paper provides a comprehensive analysis of the terms-of-trade and welfare effects of targeted export subsidies in a static three-country model of international trade. Our objective is to show the precise conditions under which a targeted export subsidy improves the welfare of the subsidizing country and the welfare effects for the subsidized and other importing and exporting countries. Using a standard general equilibrium trade model (Dixit and Norman, Bhagwati) we demonstrate that, although it is theoretically possible for targeted export subsidies to be welfare enhancing, such a welfare gain is unlikely. Similar models have been applied to agriculture by Abbott, Paarlberg, and Sharples; and Dutton.

The theoretical results shed light on important policy issues because export subsidies have become an integral part of U.S. agricultural trade policy.¹ A general equilibrium approach is taken in order to determine both the terms-of-trade and income effects associated with a targeted export

subsidy program. The income effect will be empirically important for developing countries which receive agricultural export subsidies.

We build on the transfer literature where, for example, Bhagwati, Brecher, and Hatta show that transfer paradoxes (where the donor is made better off and the recipient is immiserized) occur only when “foreign distortions” are present.² We examine the welfare impacts for the subsidizing, subsidized, and other importing and exporting countries. Precise relationships among market shares, import demand elasticities, and export supply elasticities are derived to demonstrate the conditions necessary for a targeted export subsidy program to benefit or harm the different groups of countries.

The first section of this paper is a literature review. The second section develops the general equilibrium model employed to determine the welfare impacts of targeted export subsidies on all countries. Next, the effect of the subsidy on prices is investigated. The fourth section analyzes the welfare impacts on each country. Finally, a section discusses applications to empirical research, and then a summary section concludes the paper.

Mary Bohman is an assistant professor, Department of Agricultural Economics, University of British Columbia. Colin A. Carter is a professor, Department of Agricultural Economics, University of California, Davis. Jeffrey H. Dorfman is an assistant professor, Department of Agricultural Economics, University of Georgia. Authorship is jointly shared.

Giannini Foundation Paper No. 958.

Helpful comments and suggestions have been received from Giovanni Anania, Robert Feenstra, seminar participants at the University of Southern California, and three referees.

¹ For a good description of how U.S. agricultural export subsidy programs operate, see Seitzinger and Paarlberg (1989).

² The term “foreign distortion” is taken from an earlier Bhagwati piece. If a country is not using an optimal tariff, Bhagwati has referred to this as a “foreign distortion.” While we agree with one of the referees that this is misleading terminology (because such a tariff will lower world welfare), we continue to use it to be consistent with past literature.

Previous General Equilibrium Analyses

Global export subsidies are known to be welfare reducing in general, but special cases where they can be welfare increasing have been demonstrated by Brander and Spencer, Feenstra, and Itoh and Kiyono. Abbott, Paarlberg, and Sharples (APS) have examined targeted export subsidies and found that they too can be welfare improving. As with the transfer literature, these cases can be shown to result from either domestic or "foreign distortions" (e.g., failure to exploit the optimal tariff). In the case of imperfect competition, there also exist circumstances when export subsidies are part of a first-best policy.

Itoh and Kiyono use a model with three goods to show that subsidies on commodities which are marginal goods (defined as goods not exported at all or exported in small quantities under free trade but whose exports can be promoted considerably by export subsidies) can increase the welfare of the subsidizing country. This occurs because of effects on the production structure. A subsidy on marginal goods causes their production to increase and the supply of nonmarginal goods to decrease, thereby raising the price of the nonmarginal goods and increasing the exporter's terms-of-trade. The "distortion" in place is the failure to take advantage of the optimal tariff for the nonmarginal good.

Feenstra presents a case where export subsidies increase the welfare of the subsidizing country in a three-good, two-country model. He demonstrates that it is possible for the pattern of substitutability and complementarity across the three goods to allow for subsidies to increase welfare. The necessary condition is that the subsidized export be a stronger substitute of another export good, or stronger complement of an import good, in the subsidizing country than abroad. The distortion in this case is the failure to exploit market power in the second or third good. Feenstra points out that the gain in welfare results from nonzero terms-of-trade effects in the first good. However, the subsidizing country must also be large in the good where it gains from the change in relative prices. Failure to exploit this market power creates a "foreign distortion."

The imperfect competition literature presents a case for export subsidies as a first-best policy (Brander and Spencer). This occurs because, by allowing a firm to precommit to a higher level of output and therefore worsening the terms-of-trade, profits are shifted from the foreign to the domestic firm. These results are not robust to

changes in the nature of the strategic interdependence between the two firms.

APS develop a general equilibrium model for targeted export subsidies and show that the welfare effect for the subsidizing country has an ambiguous sign. Therefore, a welfare improvement is possible as a result of a targeted subsidy. They also briefly outline some of the market conditions necessary for this result to occur. APS use a three-country model with a subsidizing country, a targeted country, and the rest of the world. However, as Dutton points out, it is not clear from the APS paper why their paradoxical result arises. Dutton shows the APS result arises because they first constrain the export tax to the rest of the world to be zero and it then becomes theoretically possible for the best export tax to the targeted country to be negative. This is a second-best price discrimination strategy. In this case, the prohibition of export taxes results in too many exports to a country with less elastic demand. Therefore, a targeted export subsidy to a country with elastic demand shifts exports out of the less elastic market. This result hinges not only on the differences in elasticities across importing countries, but also on the exporting country having market power in trade.

The General Equilibrium Model

The model employed here consists of three countries (α , β , and γ) and two goods (X and Y). Both goods are produced and consumed by each country. The export subsidy is applied to good X , and it is assumed that free trade in Y exists. Country α is always an exporter who offers an export subsidy to country γ . Country γ is always an importer with no border policies and is the recipient of the export subsidy on good X from country α . Country β , the neutral country, is either an exporter or an importer and has no border policies. Country β can be thought of as the aggregate of all the other countries. If β is composed of a group of individual countries, assumed to have no border policies, then there is no loss of generality in treating all trade flows entering or leaving the group as if they originate from or go to a single country. When country β is referred to as an importer, this is equivalent to the rest of the world being a net importer. Referring to country β as an exporter is equivalent to saying that country α 's exports are less than country γ 's imports. In this case, the market will reassign the trading flows so that α ex-

ports only to γ as a result of the targeted export subsidy. Thinking of country β as a conglomeration in this sense allows the concept of a targeted export subsidy in the three-country model to carry over to the case when β is a net exporting region.

The objective of the analysis is to determine the effect of the targeted export subsidy on the welfare levels of the three countries. The following notation is used in the model: q^i is the relative price of good X in country i ; u^i , the welfare level of country i ; S , the per unit export subsidy on good X ; $e^i(q^i, u^i)$, the expenditure function of country i ; $r^i(q^i)$, the revenue function of country i ; and $x^i(q^i, u^i)$, the compensated net import demand function for good X by country i .

The price linkages among the different countries are crucial to the results obtained. There are several cases which can occur:

(i) Country β is an importer. Country α pays a global subsidy to both β and γ . The price linkage equation is

$$(1) \quad q^\alpha = q^\beta + S = q^\gamma + S.$$

(ii) Country β is an importer. Country α pays a subsidy for exports to country γ but sells to country β at the same price that prevails in α 's domestic economy. This is a targeted export subsidy case as in APS and Dutton, where, given market power by the exporter, the optimal policy would be to price discriminate. In this case the exporter fails to fully price discriminate; therefore, the subsidy can be seen as a second-best price discrimination strategy. The standard requirement of separation of markets is assumed. The price linkage equation is

$$(2) \quad q^\alpha = q^\beta = q^\gamma + S.$$

(iii) Country β is another exporter. If country α pays a subsidy on its exports to country γ , then country β also will sell to country γ at this lower price in order to remain competitive. The price linkage equation is

$$(3) \quad q^\alpha = q^\beta + S = q^\gamma + S.$$

Note that country β does not actually make a subsidy payment to country γ . Country β sells at the prevailing world price which is defined as the price paid by the importer and not the domestic price in the exporting country.

Cases (i), (ii), and (iii) are analyzed simultaneously by writing the price linkage equation as follows:

$$(4) \quad q^\alpha = q^\beta + I(\beta)S = q^\gamma + S.$$

$I(\beta)$ is an indicator function. $I(\beta) = 1$ when the neutral country's domestic price matches the domestic price in country γ . This occurs in cases (i) and (iii). $I(\beta) = 0$ when the neutral country's domestic price equals the domestic price in the subsidizing country, α . This takes place in case (ii).

The three cases differ in their assumptions about the price linkages between different countries and correspond to different trading patterns. Case (i), the general subsidy case, results in all trade taking place at one price. Country β receives a lower price due to the subsidy. In cases (ii) and (iii), only the selected importing country γ receives the subsidy and world trade occurs at two different prices. Each individual exporting or importing country will trade only at the most favorable price available. In case (ii), the subsidizing country will export to both the subsidized and nonsubsidized importing region. Exporting countries within the nonsubsidized region will only sell to nonsubsidized importing countries. The existence of two separate markets rules out the possibility of arbitrage. In case (iii), the net exporting region (β) sells to the net importing region (γ) at the subsidized price, which is also its domestic price since it has no border measures. The subsidizing exporter (α) sells only to the subsidized region.

The model employs duality theory in terms of compensated demand functions. The format follows Bhagwati, Brecher, and Hatta. As in their model, an overspending function is defined as follows:

$$(5) \quad c^i(q^i, u^i) \equiv e^i(q^i, u^i) - r^i(q^i), \quad i = \alpha, \beta, \gamma.$$

The value of this function gives the difference between total expenditures and revenues. For the country imposing the subsidy it represents minus the cost of the subsidy. The model can be written with four equations in four unknowns: q^α , u^α , u^β , and u^γ . The equations are as follows:

$$(6) \quad c^\alpha(q^\alpha, u^\alpha) + S[x^\gamma(q^\alpha - S, u^\gamma) + I(\beta)x^\beta(q^\alpha - I(\beta)S, u^\beta)] = 0,$$

$$(7) \quad c^\beta(q^\alpha - I(\beta)S, u^\beta) = 0,$$

$$(8) \quad c^\gamma(q^\alpha - S, u^\gamma) = 0,$$

$$(9) \quad x^\alpha(q^\alpha, u^\alpha) + x^\beta(q^\alpha - I(\beta)S, u^\beta) + x^\gamma(q^\alpha - S, u^\gamma) = 0.$$

Equations (6) through (8) are the budget constraints of the three countries. The subsidy cost

in equation (6) is based on the quantity sold by α at the subsidized rate. Equation (9) is the market-clearing equation for good X. By Walras' law, the market-clearing equation for good Y can be omitted. The price linkage equation (4) can be used to solve for q^β and q^γ once q^α is known.

Comparative Statics

The price and welfare effects of changes in S in cases (i), (ii), and (iii) are analyzed through total differentiation of the model. Throughout the paper, subscripts are used to indicate partial differentiation with respect to a variable. For example, $x_q^\alpha = \partial x^\alpha / \partial q$. We assume without loss of generality that $e_u^\alpha = e_u^\beta = e_u^\gamma = 1$ initially. In addition, $S = 0$ initially. Using Shepard's lemma we know that $c_q^i = x^i$, where x^i is the net import demand of country i . Finally, define $x_q \equiv x_q^\alpha + x_q^\beta + x_q^\gamma$. The differentiated system can be written as follows:

$$(10) \quad \begin{bmatrix} x^\alpha & 1 & 0 & 0 \\ x^\beta & 0 & 1 & 0 \\ x^\gamma & 0 & 0 & 1 \\ x_q & x_u^\alpha & x_u^\beta & x_u^\gamma \end{bmatrix} \begin{bmatrix} dq^\alpha \\ du^\alpha \\ du^\beta \\ du^\gamma \end{bmatrix} = \begin{bmatrix} -I(\beta)x^\beta - x^\gamma \\ I(\beta)x^\beta \\ x^\gamma \\ x_q^\gamma + I(\beta)x_q^\beta \end{bmatrix} dS.$$

Applying Cramer's rule to this system we obtain the following results, where $\Delta \equiv -(x_q - x^\alpha x_u^\alpha - x^\beta x_u^\beta - x^\gamma x_u^\gamma)$ and $\bar{x}_q^i = (x_q^i - x_u^i x^i)$ is the partial derivative of the Marshallian import demand function with respect to the relative price.

$$(11) \quad dq^\alpha/dS = -\{\bar{x}_q^\alpha + I(\beta)\bar{x}_q^\beta + x_u^\alpha(x^\gamma + I(\beta)x^\beta)\}/\Delta,$$

$$(12) \quad du^\alpha/dS = -\{x^\gamma \Delta + x^\alpha[\bar{x}_q^\gamma + I(\beta)\bar{x}_q^\beta + x_u^\alpha(x^\gamma + I(\beta)x^\beta)]\}/\Delta,$$

$$(13) \quad du^\beta/dS = \{x^\beta[I(\beta)\Delta - \bar{x}_q^\gamma - I(\beta)\bar{x}_q^\beta - x_u^\alpha(x^\gamma + I(\beta)x^\beta)]\}/\Delta,$$

$$(14) \quad du^\gamma/dS = \{x^\gamma[\Delta - \bar{x}_q^\gamma - I(\beta)\bar{x}_q^\beta - x_u^\alpha(x^\gamma + I(\beta)x^\beta)]\}/\Delta.$$

The denominator of each expression is the determinant of the (4×4) matrix in (10). This determinant, Δ , equals minus the slope of the

Marshallian general equilibrium global excess-demand schedule of good X. If we assume that the system is stable, then the Marshall-Lerner condition implies that $\Delta > 0$.

The Effect on Prices

In a partial equilibrium model, all results are caused by price (or terms-of-trade) effects. In contrast, the results of a general equilibrium model are also influenced by income effects. It is useful to analyze the price (terms-of-trade) effects separately. This clarifies the role that changes in prices and the terms-of-trade play in determining the welfare effects of targeted, and general, subsidies. At the same time, we have to remind ourselves that this is only part of the story.

The export subsidy's effect on prices can be investigated using equation (11). It is important not to interpret the domestic price as the world price because there is no case where all world trade takes place at q^α . In case (i), with a general subsidy and case (iii), with two exporters, all trade (although not all domestic sales) occurs at $q^\alpha - S$. In case (ii), with a targeted subsidy and one exporter, only a portion of trade takes place at q^α .

Case (ii)

In this case, β is an importer and α offers a targeted export subsidy to γ . The relative magnitude of the Marshallian income effects in countries α and γ determines whether dq^α/dS is positive. To see this, rewrite (11) as

$$(15) \quad dq^\alpha/dS = -\{x_q^\gamma - x^\gamma(x_u^\gamma - x_u^\alpha) + I(\beta)[x_q^\beta - x^\beta(x_u^\beta - x_u^\alpha)]\}/\Delta.$$

Recalling that in case (ii) $I(\beta) = 0$, a sufficient condition for dq^α/dS to be positive is that x_u^γ is greater than or equal to x_u^α . Note that because e_u^i is defined to equal 1 at the point in question, x_u^i equals the Marshallian income effect. Intuitively, $dq^\alpha/dS > 0$ requires that the increase in demand in the subsidized country exceeds the reduction in demand in the subsidizing country as a result of the changes in each country's domestic price.

Cases (i) and (iii)

In case (iii), the targeted export subsidy with two exporters, the sufficient condition for dq^α/dS

dS to be positive is that $x_u^\gamma > x_u^\alpha > x_u^\beta$. This can be seen in equation (15), remembering that x^β is negative for this case. In addition, when β is an exporter the value of dq^α/dS is always less than one, since (11) can be rewritten as

$$(16) \quad dq^\alpha/dS = 1 + x_q^\alpha/\Delta.$$

Continuing with case (iii), the effect of the subsidy on prices in α is more positive as α 's share of the export market declines. This results from the fact that as α becomes more of a price taker, its actions have less impact on the international market. Algebraically, this can be seen by expressing dq^α/dS in terms of elasticities and market shares. First, define ϵ^α as the price elasticity of demand for country α 's exports and ϵ^w as the price elasticity of the world excess demand function for the case where only country γ is an importer ($\epsilon^\alpha = x_q^\alpha(q/x^\alpha)$ and $\epsilon^w = \Delta(q/x^\gamma)$).³ Further, let κ be country α 's share of the export market; with β an exporter, this implies that $\kappa = x^\alpha/(x^\alpha + x^\beta)$. The necessary and sufficient condition for dq^α/dS to be positive is that $\epsilon^\alpha < \epsilon^w/\kappa$.⁴ As country β exports more, the right-hand side of the elasticity inequality becomes larger (as κ becomes smaller) and the condition is more easily satisfied (since both elasticities are defined to be positive).

In case (i), the global subsidy, a sufficient condition for dq^α/dS to be positive is that both x_u^α and x_u^β are greater than x_u^γ . The condition on the magnitudes of the income effects in α and β switches direction in case (iii) because β is an importer.

Welfare Effects

To analyze the welfare effects that are caused by the subsidy, equations (12) through (14) are rewritten below in terms of dq^α/dS and the quantities traded.

³ Note that a single price, q , can be used here since all expressions are being evaluated around an initial subsidy level of zero.

⁴ The proof of this result is as follows:

From eq. (16): $(dq^\alpha/dS) = 1 + x_q^\alpha/\Delta$.

Thus, we want

$$\begin{aligned} (dq^\alpha/dS) &= 1 + x_q^\alpha/\Delta > 0, \\ 1 + x_q^\alpha/\Delta &> 0, \\ x_q^\alpha/\Delta &> -1, \\ x_q^\alpha &> -\Delta, \\ x_q^\alpha &> -\Delta q, \\ x_q^\alpha(q/x^\alpha) &< -\Delta(q/x^\alpha), \\ \epsilon^\alpha &< -\Delta(q/x^\alpha), \\ \epsilon^\alpha &< -\Delta(q/x^\gamma)(x^\gamma/x^\alpha), \\ \epsilon^\alpha &< -\epsilon^w(x^\gamma/x^\alpha), \\ \epsilon^\alpha &< -\epsilon^w[(x^\alpha - x^\beta)/x^\alpha], \\ \epsilon^\alpha &< -\epsilon^w/(x^\alpha/[-x^\alpha - x^\beta]), \\ \epsilon^\alpha &< \epsilon^w/\kappa. \end{aligned}$$

$$(17) \quad du^\alpha/dS = -(x^\gamma + I(\beta)x^\beta) - x^\alpha(dq^\alpha/dS),$$

$$(18) \quad du^\beta/dS = x^\beta\{I(\beta) - (dq^\alpha/dS)\} \\ = -x^\beta(dq^\beta/dS),$$

$$(19) \quad du^\gamma/dS = x^\gamma(1 - dq^\alpha/dS) \\ = -x^\gamma(dq^\gamma/dS).$$

These equations highlight the role of the change in the terms-of-trade. The welfare result in equation (17) is composed of an income effect and a terms-of-trade effect. In equations (18) and (19) the welfare effect depends entirely on the terms-of-trade effect because these countries incur no subsidy costs.

Effects on the Subsidizing Country

The welfare of the subsidizing country always falls when the neutral country is an exporter. However, the welfare effect is ambiguous with a targeted subsidy when the neutral country is an importer.

Cases i and iii. In cases (i), a global subsidy, and (iii), where the neutral country is an exporter, $-(x^\gamma + I(\beta)x^\beta)$ equals x^α . For these two cases, the welfare loss can be seen by simplifying (17) when $I(\beta) = 1$ to

$$(20) \quad du^\alpha/dS = x^\alpha(1 - dq^\alpha/dS).$$

Therefore, utilizing equation (16) which shows that dq^α/dS is less than one when $I(\beta) = 1$, du^α/dS is unambiguously negative with a targeted subsidy when country β is an exporter and for a global subsidy. Thus, we conclude that for both cases (i) and (iii), $du^\alpha/dS < 0$. As in the case of the two-country model of standard trade theory, offering a general subsidy is always welfare decreasing in the three-country case both when offered to all countries and when a subsidy is targeted to one importing country in the presence of another exporter or net exporting group.

Case ii. In case (ii), with the neutral country an importer, the possibility arises of a paradox where the subsidy causes welfare in the subsidizing country to rise. Using equation (17) the necessary and sufficient condition for du^α/dS to be positive when the third country is an importer is that $dq^\alpha/dS > \lambda$ where λ is γ 's share of world imports, $\lambda = x^\gamma/(x^\beta + x^\gamma)$. Thus, as the share of α 's exports to γ increases (and, conse-

quently, the share to country β decreases), du^α/dS tends to be negative because λ increases. This occurs because country α is exploiting β through the targeted subsidy to γ . The larger β 's share of country α 's exports, the more opportunity to earn extra revenue by the price discriminating effects of a targeted subsidy, and the more likely that country α 's subsidy will be welfare improving.

The ambiguous sign of du^α/dS for case (ii) can be explained by the fact that the exporter is failing to take advantage of its market power in good X . The use of the subsidy captures some of the benefits of α 's power in β 's and γ 's markets but still leaves country α with less than the maximum attainable welfare. In a similar manner to the case of transfers, the paradox arises from the presence of a "distortion" (the failure to exercise market power). See Dutton for further clarification of this result.

Effects on the Neutral Country

The expression in (18) for country β 's welfare change can be interpreted solely as a terms-of-trade effect. The welfare of the neutral country falls when the change in its terms-of-trade is unfavorable.

Cases i and iii. Country α 's decision to offer a subsidy reduces not only its own welfare but also that of all other exporters. The effect of the subsidy depends on whether the neutral country is an exporter or an importer. Remembering that $I(\beta) = 1$ for both cases (i) and (iii), in case (i), with a global subsidy, β 's welfare improves. In case (iii), with a targeted subsidy when β is an exporter, $dq^\alpha/dS < 1$ by equation (16), and β must suffer a welfare loss as a result of the subsidy offered by α . The change in sign of the welfare impact on country β is caused by the switch in sign of x^β , the net import demand of country β , between the global subsidy, case (i), where β is an importer and case (iii) where β is an exporter. Country α 's decision to offer a subsidy reduces not only its own welfare, but also that of all other exporters.

Case ii. In case (ii) country β is an importer not receiving α 's targeted export subsidy. By equation (18) du^β/dS equals minus net imports times the change in the price of its imports due to the targeted subsidy. In this case the price in

β is equal to the price in α , therefore country β is made worse off if and only if $dq^\alpha/dS > 0$. If country β is made better off, country α must suffer a loss in welfare. The exact necessary and sufficient condition for $dq^\alpha/dS < 0$ can now be expressed in terms of elasticities to help in a later comparison of the welfare changes in β and γ . Thus, for case (ii), $du^\beta/dS > 0$ if and only if

$$(21) \quad q(x_u^\alpha - x_u^\gamma) > \epsilon^\gamma.$$

In (21), ϵ^γ is the (positively defined) elasticity of demand for good X in country γ , $\epsilon^\gamma = -x_q^\gamma(q/x^\gamma)$. This specification agrees with the sufficient condition for dq^α/dS concerning the direction of the price change ($x_u^\alpha > x_u^\gamma$) but extends it to the necessary and sufficient condition. Thus, as an importer, country β suffers a drop in utility when the domestic price in country α rises.

When it is an importer, β is the country most affected by the difference between a targeted or a general subsidy from α . Country β would benefit from a global subsidy which would lower the price of its imported good but will normally lose if α only offers a targeted export subsidy to γ .

Effects on the Targeted Country

In general country γ 's welfare increases as a result of the subsidy. Only in case (ii) when β is an importer is the paradoxical result that γ is immiserized by the targeted subsidy from α possible. As in the case of country β , the welfare change in country γ is a straightforward terms-of-trade effect. Irrespective of whether β is an importer or an exporter du^γ/dS is positive if and only if dq^γ/dS is negative, representing a terms-of-trade gain.

Cases i and iii. In case (iii) when country β is an exporter, country γ 's welfare improves with a subsidy. Equation (16) shows that $[1 - (\epsilon q^\alpha/dS)]$ must be positive for case (iii) which ensures that γ is made better off. As in the case of subsidies in the classical two-country model of trade, in the presence of another exporter, the importing country cannot be immiserized by the subsidy. The result also holds for case (i), the global subsidy.

Case ii. In this case, when the neutral country is an importer, the possibility of a paradoxical

result exists. If dq^α/dS is greater than unity, then country γ is immiserized. Returning to the earlier elasticity representations, the necessary and sufficient condition for this to occur when β is an importer is⁵

$$(22) \quad q(x_u^\gamma - x_u^\alpha) > (\epsilon^w/\lambda) - \epsilon^\gamma.$$

The price elasticity of the Marshallian world excess demand (defined to be positive) is given by $\epsilon^w = \Delta[q/(x^\beta + x^\gamma)]$ to reflect the fact that country β is an importer, $\epsilon^\gamma = -x_q^\gamma(q/x^\gamma)$ as before, and λ is, as above, the share of total world imports imported by the targeted country, $\lambda = x^\gamma/(x^\gamma + x^\beta)$. The paradoxical result of γ being made worse off by a subsidy can occur if and only if there exists a large difference in the income effects of countries α and γ .

The condition in (22) can be rewritten as follows to offer additional intuition.

$$(23) \quad x_q^\beta + x_q^\gamma + x^\beta(x_u^\alpha - x_u^\beta) > 0.$$

From (23) it can be seen that a large income effect in α will make the paradox more likely. The result is similar to the Metzler paradox. Recall that the x functions are net import demand functions. The subsidy initially results in the import price falling in γ and a rising in β . If the rise in the import price has a small impact on β 's net import demand (due to small price and income effects), then α 's exports to β will not fall much. This condition, combined with the increased (subsidized) exports to γ , causes domestic supply in α to fall and increases the domestic price. If the domestic price increases enough, even the subsidized price will be above the pre-subsidy price level; this scenario results in the immiserization of country γ .

The conditions necessary for the paradoxical result that γ is made worse off precludes the subsidy improving β 's welfare. A comparison of (21) and (22) shows that in case (ii) when country β is an importer, the condition for β to be made better off from being (price) discriminated against (i.e., not receiving a subsidy) can be met if and only if the subsidy is welfare increasing for country γ . This result follows from the fact that α can only gain utility if $dq^\alpha/dS > \lambda$, which is positive by definition. Equation (18) shows that if $dq^\alpha/dS > 0$, then du^β/dS is negative when β is an importer. It can also be shown, using (17) and (19), that if α loses utility from the subsidy, then γ must gain utility. This oc-

curs because when α loses utility $dq^\alpha/dS < \lambda < 1$, and from (19) this implies that γ must gain utility. Therefore, in the unlikely, but possible, event that γ suffers a loss in utility by receiving a subsidy, α will gain utility and β will lose utility. The above comparison of du^α/dS to λ allows a simple proof by contradiction of this last result.

Summary of Comparative Static Results

The price and welfare effects of the three cases above are summarized in table 1. Case (i) represents a global subsidy to both the neutral and subsidized country; case (ii) represents a targeted subsidy to the subsidized country, with the neutral country as an importer; case (iii) represents a targeted subsidy to the subsidized country with the neutral country as an exporter. The price effects are summarized in the top panel of table 1 and the welfare effects in the bottom panel.

Implications for Policy and Empirical Research

The results from the general equilibrium model can be used to evaluate recent empirical research on export subsidies for U.S. agricultural products. Empirical applications have generally been partial equilibrium in nature and most have focused on the U.S. Export Enhancement Program for wheat. In general, welfare calculations are not made. This section discusses the conclusions from the theoretical model that apply to policy issues and empirical research and evaluates some recent models.

Changes in the terms-of-trade can be used to infer welfare results for the different groups of countries. Equations (17)–(19) show that the changes in welfare hinge on changes in the terms-of-trade for each country. The empirical models often calculate changes in the terms-of-trade but report only changes in total revenue and not price changes. Given the importance of the terms-of-trade to welfare changes, future research should report this result.

The theoretical results also show the importance of the share of the market which is subsidized. The λ and κ parameters representing market share appear in the welfare conditions summarized in table 1. Recall that λ is γ 's share of world imports when β is an importer, and κ is α 's share of the export market with β an ex-

⁵ Because $S = 0$, $q^\alpha = q^\beta = q^\gamma = q$. We need only q^i for differentiation, e.g., dq^α/dS .

Table 1. Summary of Welfare Results

Country	Case <i>i</i>	Case <i>ii</i>	Case <i>iii</i>
		Price Effects $dq^a/dS > 0$ if	
Exporter(α)	$x_u^\gamma > x_u^\beta > x_u^\alpha$	$x_u^\gamma > x_u^\epsilon$	$\epsilon^\alpha < \epsilon^w/\kappa$
		Welfare Effects $du^i/dS > 0$ if	
Exporter (α)	Never	$dq^a/dS > \lambda$	Never
Neutral (β)	Always	$*q(x_u^\alpha - x_u^\gamma) > \epsilon^\gamma$	Never
Subsidized (γ)	Always	$*q(x_u^\gamma - x_u^\alpha) > \frac{\epsilon^w}{\lambda} - \epsilon^\gamma$	Always

Notes: Case *i*: global subsidy to both the neutral (β) and subsidized (γ) countries; Case *ii*: targeted subsidy to γ , with β an importer; Case *iii*: targeted subsidy to γ , with β an exporter. x_u^i is Marshallian income effect; ϵ^α is price elasticity of demand for country α 's exports; ϵ^w is price elasticity of demand of the world excess demand function; κ is exporter's (α 's) share of the export market for case *iii*; λ is subsidized country's (γ 's) share of imports for case *ii*; and ϵ^γ is absolute value of price elasticity of demand in subsidized country (γ). Asterisk indicates it cannot happen simultaneously.

porter. In the case of the EEP, the percentage of U.S. exports receiving subsidies has grown over time to the point where in 1988, EEP sales approached 29 million metric tons (mmt), which was 70% of U.S. wheat sales. The results of any model designed to analyze the program are sensitive to the time period chosen. Therefore, the possibility of an increase in U.S. welfare from the EEP has fallen over time and the results in Seitzinger and Paarlberg (1990), for example, are probably no longer valid.

Results from the theoretical model may not be directly comparable to those from various empirical models because of different underlying assumptions. The theoretical model presented in this paper assumes perfect competition and no other policy intervention besides the export subsidy. The price linkage equations reflect these assumptions and result in specific trade flows. The different trade flows that have occurred with EEP could reflect the fact that the world wheat market is not perfectly competitive. Also, the theoretical model does not incorporate policies other than export subsidies. Results from empirical models which contradict the theoretical welfare effects could occur because of the way that stocks, support prices and other policies are modeled. General welfare effects cannot be derived if other distortions exist.

APS used a spatial equilibrium model to calculate optimal targeted export subsidies. They found that small subsidies resulted in a small welfare gain relative to free trade for the subsidizing country. This is consistent with the theoretical result that welfare-improving targeted subsidies occur under constrained price discrimination in our case *ii*.

Seitzinger and Paarlberg (1990) analyze quarterly EEP sales using a game theory approach and a spatial equilibrium model. They do not calculate overall welfare effects. However, they do calculate changes in the terms-of-trade for the United States and changes in total revenue for other exporters. They estimate that the U.S. border price is higher with EEP for October to December 1985. This result could only be supported by our theoretical results if the share of EEP sales in total U.S. exports was small.

Haley used a spatial equilibrium model to analyze the effect of the EEP in 1986/87 but did not calculate welfare changes. In Haley's model, subsidies exist with shared markets between the subsidizing exporter and nonsubsidizing exporters. He found that subsidies lowered the prices received by competing exporters which agrees with our theoretical result. Similarly, to Seitzinger and Paarlberg (1990), he finds the U.S. export price rises, which is possible but not probable, given our theoretical results.

Conclusions

This paper analyzes the welfare impacts of targeted export subsidies using a three-country, two-good model. The commercial policies of each country and its position in the world market (exporter or importer) are exogenous to the model. One country subsidizes its exports, and there is a recipient and a neutral country. The results show that the neutral country's market position (importer or exporter) and its market share influence the welfare results for all countries. The effect of the subsidy on the domestic price in

the country offering the subsidy, the slopes of the various excess demand functions, and the magnitudes of the income effects are all important to the welfare results.

When the neutral country is an exporter, the subsidizing country must suffer a loss of welfare as a result of offering a targeted subsidy. Further, as an exporting country, the neutral country must suffer a loss in welfare while the subsidized country benefits from the subsidy. With two importing countries, a targeted export subsidy reduces the welfare of the subsidizer unless targeted toward a country with a small market share and/or a higher than normal income elasticity of demand.

In general, the qualitative effects of a targeted export subsidy differ little from a general subsidy. A targeted subsidy imposes less of a welfare cost because it is offered to only part of the market. However, if an exporter is in a position such that it can gain through offering a targeted subsidy, then it should take full advantage of its market power and price discriminate using export taxes if it chooses to maximize its welfare by abandoning free trade to a beggar-thy-neighbor policy (see Dutton).

Export subsidies have been shown to be similar to (international aid) transfers. As demonstrated in Bhagwati, Brecher, and Hatta, a three-country model is essential to finding the paradoxical results where the country being granted aid in the form of a transfer or subsidy is actually immiserized by it. The neutral third country is crucial because this paper has shown that such a paradox can exist only when the third country is an importer and is not receiving the subsidy. Similarly, in the Bhagwati, Brecher, and Hatta case, a transfer can only immiserize the recipient country when the neutral country is an exporter. The case of targeted export subsidies also mirrors the case of transfers in that only one of the two importers benefits from the subsidy while the other loses welfare. Thus, when the subsidized country is immiserized, the neutral country benefits. In a final parallel to the transfer case, it is possible for the neutral country to lose welfare while the subsidized and subsidizing country both gain welfare. This case is an exact analogy to what Bhagwati, Brecher, and Hatta refer to as the "invisible shakedown."

Targeted export subsidies for agricultural exports have been used extensively by the United States since 1985 in order to regain market share. However, the model presented in this paper fails to provide justification (on national welfare grounds) for the widespread use of targeted ex-

port subsidies. Instead, the results of this paper suggest that a targeted export subsidy program is likely to be suboptimal trade policy. In agricultural trade, the neutral countries tend to be net exporters, and thus the United States stands to lose welfare as a result of targeted export subsidies.

The result that the welfare of the subsidizing country falls as the share of the subsidized market increases argues against expansion of current programs. The growth of export subsidy programs over the past few years has been at the expense of national welfare. One caveat is the use of subsidies as part of strategic trade policy to induce other exporters to change their trade policies. This dynamic, multiperiod strategy is not incorporated into the static model used in this paper.

This paper also demonstrates why exporting neutral countries have objected so strongly to subsidy programs. The EEP was originally designed to punish the EC for their export subsidies (as explicitly modeled by Seitzinger and Paarlberg). The U.S. Department of Agriculture claimed that other exporters would not be damaged. However, Canada and Australia immediately recognized that U.S. export subsidies adversely affected their terms of trade. Their complaints are supported by the results of this model.

[Received October 1989; final revision
received October 1990.]

References

- Abbott, P., P. Paarlberg and J. Sharples. "Targeted Export Subsidies and Social Welfare." *Amer. J. Agr. Econ.* 69(1987):723-32.
- Bhagwati, J. N. "The Generalized Theory of Distortions and Welfare." *Trade, Balance of Payments and Growth*, ed. J. N. Bhagwati et al. Amsterdam: North-Holland Publishing Co., 1971.
- Bhagwati, J. N., R. A. Brecher, and T. Hatta. "The Generalized Theory of Transfers and Welfare: Bilateral Transfers in a Multilateral World." *Amer. Econ. Rev.* 73(1983):607-18.
- Brander, J. A., and B. J. Spencer. "Export Subsidies and International Market Share Rivalry." *J. Int. Econ.* 18(1985):83-100.
- Dixit, A. K., and V. Norman. *Theory of International Trade*. Digswell Place, Welwyn Herts: James Nisbet and Co. and the Cambridge University Press, 1980.
- Dutton, J. "Targeted Export Subsidies as an Exercise of Monopoly Power." *Can. J. Econ.* 23(1990):705-10.
- Feenstra, Robert C. "Trade Policy with Several Goods and 'Market Linkages.'" *J. Int. Econ.* 20(1986):249-67.

- Haley, Stephen L. *Evaluation of Export Enhancement, Dollar Depreciation and Loan Rate Reduction for Wheat*. Washington DC: U.S. Department of Agriculture, Econ. Res. Serv., Agr. and Trade Anal. Div. Staff Rep. No. AGES 89-6, 1989.
- Itoh, M., and K. Kiyono. "Welfare-Enhancing Export Subsidies." *J. Polit. Econ.* 95(1987):115-37.
- Seitzinger, A. H., and P. L. Paarlberg. "A Simulation Model of the U.S. Export Enhancement Program for Wheat." *Amer. J. Agr. Econ.* 72(1990):95-103.
- . *A Survey of Theoretical and Empirical Literature Related to Export Assistance*. Washington DC: U.S. Department of Agriculture, Econ. Res. Serv. Commodity Econ. Div. Staff Rep. No. AGES 89-34, 1989.

The Costs of Indonesian Sugar Policy: A Policy Analysis Matrix Approach

Gerald C. Nelson and Martin Panggabean

Indonesian sugar policy is a complex web of contradictory policies, including mandatory production, price supports, and fertilizer and credit subsidies. The policy analysis matrix (PAM) was developed by Monke and Pearson to provide a more complete perspective on social profitability and the divergence between private and social costs than other commonly used social cost-benefit measures. The PAM is used to analyze the effects of Indonesian sugar policy on sugar production in irrigated and dryland areas on Java, the main sugar-producing region in Indonesia, and to identify the distribution of resource transfers.

Key words: food policy, Indonesia, policy analysis matrix, social cost-benefit analysis, sugar.

Policy analysts who use the tools of applied welfare economics face the challenge of presenting their results to policy makers who have little time to spare and varying degrees of analytical sophistication and training in welfare economics. Hence, a central part of theoretical efforts in this field has been the search for theoretically correct and easily understandable summary measures of policy impacts on social welfare. Two types of summary measures have been developed. One strand of analysis focuses on the private and social costs of public sector investment. Popular measures in this area include the net present value (NPV), and the economic internal rate of the return (EIRR). (See Gittinger for a widely used reference to these measures.) The second strand of analysis focuses on the static effects of price-distorting policies. Popular measures of the effects of price policies include the effective protection coefficient (EPC) (e.g., Corden, Balassa and Schydowsky) and the domestic resource cost of foreign exchange (DRC). (See Krueger, and Bruno 1967 for the original statement, a series of ar-

ticles in the *Journal of Political Economy* in the early 1970s, e.g., Bruno 1972, and Pearson, Akrasanee, and Nelson for a restatement and use in an agricultural context.)

Each of these measures has advantages and disadvantages, depending upon the type of analysis conducted. However, the danger with any summary measure is that it may summarize too much, and important results of the analysis go unnoticed. A new technique, the policy analysis matrix (PAM) addresses this problem (Monke and Pearson). The PAM was developed by Pearson et al. and grew out of a history of policy analysis using the DRC approach (e.g., Pearson and Cownie, Pearson et al., Falcon et al.). The basis of the PAM is a set of profit and loss identities that are familiar to any businessman. One strength of the PAM is that it allows varying levels of disaggregation. Another is that it makes the analysis of policy-induced transfers straightforward. Finally, the PAM makes it possible to identify the net effect of a set of complex and contradictory policies and to sort out the individual effects of those policies. As the intellectual successor to the DRC and NPV calculations, the PAM also suffers from some of their weaknesses. Principal among these is the assumption of fixed input-output coefficients (although Monke and Pearson argue that it is possible to introduce supply response parameters into the PAM). Production is characterized by a series of techniques, each of which has fixed

The authors are an assistant professor and a doctoral candidate, respectively, Department of Agricultural Economics, University of Illinois.

They would like to thank Sam H. Johnson for comments on an earlier draft, Laurian J. Unnevehr for substantive and editorial assistance throughout the process of preparing this paper, and three anonymous reviewers. As always, the standard disclaimer about responsibility applies.

input-output coefficients and represents some share of the total production.

Analysis of the Indonesian sugar sector is an ideal candidate for use of the PAM. While Indonesian policy makers have a deserved reputation for managing their economy well, the Indonesian sugar sector is a conspicuous exception to that record of good economic policy. This paper uses the PAM to examine the social cost of achieving (at least through the late 1980s) the policy goal of self-sufficiency in sugar production. Four sets of policy instruments contribute to that social cost.

In the next section, a brief history of Indonesian sugar cultivation and description of current policies is presented. The third section provides a summary of the policy analysis matrix approach and discussion of data and modeling assumptions. The fourth section presents results and the final section some concluding remarks.

A Brief History of Sugar Cultivation, Trade, and Policies in Indonesia

Sugar is a processed form of sugarcane (*Saccharum officinarum*), a plant that originated in New Guinea.¹ The dispersion of sugarcane from New Guinea followed three routes, one of which passed through Indonesia. It seems likely that sugarcane has been grown in Indonesia at least since 1 A.D. (Husz, Blackburn).

Before the seventeenth century, sugar exports were small compared with spice exports, the primary reason traders came to Indonesia. The importance of sugar in Indonesia changed in 1830 when the Dutch colonists introduced the *Cultuur Stelsel* (translated freely as the Cultivation System) which was so successful that it became the mainstay of the Netherlands' prosperity (see for example, Sievers for a description of the Cultivation System). The central element of the Cultivation System was that farmers were forced to provide land and labor to produce export crops. Production decisions were made by the colonialists and their employees. In 1870, the Dutch parliament passed the Sugar Act, which removed sugar from the Cultivation System, and the Agrarian Act, which institutionalized the practice of long-term leases of land for cultivation of crops for export. The "Glebagan" sys-

tem arose from these changes. Under this system, farmers near a sugar mill were required to lease their land to the mill for one year in three.

By 1930, Indonesia was the world's second biggest sugar producer and exporter after Cuba. The 1930 Depression began the decline in the Indonesian sugar industry as the bottom dropped out of world demand. Efforts to rejuvenate production after the Depression were halted by World War II. The Japanese occupiers destroyed many sugar mills, and the Indonesian war of independence (1945-49) also saw more destruction of sugar mills.

Although the Glebagan system remained in effect after the war, the need to increase food production, especially rice, after independence made it difficult for sugar mills to continue the prewar practice of managing cane production on land leased from smallholders (Mubyarto 1969). To assure an adequate supply of cane, mills began to rely on independently grown smallholders' cane. For the first time, smallholders established an important role as cane suppliers.

Beginning in 1952, a series of legislative reforms were passed in an attempt to remove some of the inequities associated with the long-term lease provisions of the colonial Agrarian Law (Soetrisno, Mubyarto 1969). These culminated in the TRI system, described below, in 1975.

The first half of the 1970s saw a sizable recovery in cane production with average annual sugar output of 840,000 metric tons, almost 200,000 metric tons larger than production for most of the 1960s (table 1). This growth was due to two factors (Mubyarto 1977), an increase in yields and area expansion induced by movements in sugar and rice prices and an increase in areas designated for sugar production. But an important source of these improvements lies in the change of government and the improved economic environment after 1965, which led to improvement in roads, other infrastructure, administration, and management efficiency.

The compulsory leasing of farmland to mills, a crucial element of the Glebagan system, was eliminated in February 1975, when the government issued Presidential Instruction No. 9. This new law was designed to encourage farmers, rather than mills, to manage the sugar production on their land. The new program, called the Smallholder Cane Intensification Program and known by its Indonesian acronym, TRI, involved a package of credit and inputs. Instead of leasing land to the mills, farmers were forced to cultivate cane periodically to ensure an adequate supply of cane to the mills.

¹In Indonesia, the word "sugar" is used to refer to various kinds of sweeteners, some of which came from plants other than sugarcane. In this paper, however, sugar means a refined, white sugar processed from sugarcane.

Table 1. Area, Production, and Yield of Sugar on Java (Average of Selected Years)

Year	Area (000 ha)	Production (000 ton)	Sugar Yield (mt/ha)	(off-Java area) (000 ha)
1954-59	63	697	11.1	NA
1960-69	81	644	8.0	NA
1970-74	92	859	9.3	NA
1975-79	138	1,118	8.1	NA
1980-84	244	1,485	6.1	24.4
1985-87	310	1,959	6.3	66.4

Source: Mubyarto (1969, 1977) for years before 1970. After 1970 from Balai (later Pusat) Penyelidikan Perusahaan Perkebunan Gula, *Laporan Tahunan*, various years. Includes only cane grown for granular sugar production.

The main features of the TRI system, which remains in place today, are as follows: (a) Farmers whose land is in a designated cane production area are required to cultivate cane, but in not more than two consecutive seasons. After that period, they are allowed to cultivate other crops until their land is again designated for cane production, typically after two years.² (b) Farmers are provided with credit at subsidized interest rates (12%) by Bank Rakyat Indonesia (BRI). This credit is given in four tranches according to the stages of cultivation and is for both inputs and living allowances. (c) Farmers are required to use seeds, pesticides, and fertilizers distributed by the mills. (d) Farmers are required to submit all of their cane production to the mills, and the mills are required to accept and process it. (e) Payment is determined by the rate of extraction of sugar from the cane (for the typical extraction rate of 8%, farmers receive 55% of sugar output). All of the farmers' share is bought by the government at prices announced at the beginning of the planting season. Farmers receive payment in cash net of loan repayment.

After the TRI program was implemented, cane cultivation expanded to new areas not previously utilized for cane production. Cane grown on irrigated land grew at an average annual rate of 8.52% during the 1980-86 period, while dryland cane area (including land off-Java) grew at an annual average of 12.30%. Since 1984, dryland cane area has been larger than irrigated cane area (*Tim Survai* [Anonymous]; see also Brown for a more extended description of efforts to expand sugar production off-Java).

²As a grass, new cane can sprout from roots left in the ground after the first crop is cut. A typical rotation might be to plant cane, harvest the first crop 16 months after planting, allow new cane to grow from the old crop roots and harvest the second, "ratoon" crop 12 months after the first harvest. Sugar yields typically decline markedly after the second crop.

Sugar Marketing and Price Control

The history of sugar marketing in Indonesia is checkered by failed experiments with different forms of state/private management. The first collective sugar-marketing body, NIVAS (*Nederlandsch Indische Vereniging voor de Afzet van Suiker*), was established in 1931 by all of the sugar mills to secure overseas markets during the world depression (Mubyarto 1969). NIVAS' other functions were securing the collection of duties imposed on sugar and control of domestic consumption to maintain export levels (Soetrisno).

When the sugar mills were nationalized in 1957, NIVAS stopped functioning. As with production regulations, government policies with respect to processing underwent many changes during the 1960s and early 1970s. By the early 1970s, the government agencies in charge of processing and marketing were (Soetrisno): (a) Ministry of Agriculture, whose responsibility was to manage the state-owned sugar mills' production; (b) Bank Bumi Daya, which acted as the mills' creditor and sugar stockholder and was responsible for financing; and (c) BULOG (the National Logistics Agency), which was responsible for procurement, distribution, and marketing of all sugar.

In 1980, the government issued a decree stating that beginning in 1981 all domestic sugar production, regardless of ownership or method of production, would be bought by BULOG except 1% (later increased to 2%) of TRI farmers' production for use in their own consumption.

Since independence, sugar has been one of nine basic commodities whose prices are controlled. Sugar price control is implemented at the mill gate. Since 1969, the Provenue price, which is effectively the net payment received by mills, has been fixed annually by the government, and the ex-mill price is determined by

adding various fees, taxes, and miscellaneous charges. The Provenue price is determined in consultations among the Ministries of Agriculture, Trade, Finance, and BULOG and is based on the mills' cost of production plus a profit margin and an allowance for inflation. Under normal circumstances, the Provenue price for a particular year is announced at the start of the harvest season of that year's cane, which is also the beginning of the next year's cane planting season.

In recent years the world price for sugar has been much lower and more variable than the domestic price (table 2). In the 1980s, the nominal protection coefficient (NPC) has generally been between 40 and 80, with the fluctuations primarily as a result of changes in the world price.

Present policies on marketing and price control are implemented by BULOG. BULOG is the sole buyer of all domestically produced sugar, except for the 2% of farmers' output retained for own consumption. BULOG is also the sole importer of sugar. In addition, BULOG heavily influences domestic distribution of sugar (Sapuan et al.) through its role in appointing regional wholesalers who are permitted to obtain supplies from its regional warehouses.

In summary, four policy instruments affect the sugar sector directly. As in many LDCs, both producer and consumer prices of sugar are generally well above world prices.³ Second, in spite of high prices, the government has to compel some farmers on Java to plant sugarcane. Java

has unique physical characteristics that make its land suitable not only for growing sugar cane but also for producing rice and other food crops such as corn and soybeans. The fact that rice is an alternative crop presents policy makers with a special dilemma since rice self-sufficiency is the single most important agricultural policy goal. Third, fertilizers and other chemical inputs are subsidized. Finally, most cane producers receive a substantial credit subsidy.

The Policy Analysis Matrix (PAM)

The PAM consists of two accounting identities (table 3). The first identity states that profit is equal to revenue minus costs, measured in either private or social terms; the second identity measures the differences between observed values and the levels that would exist if the divergences between private and social prices (caused by distorting policies or market failures) were removed.

On the first line of the matrix is the calculation of private profitability (D), defined as revenue (A) minus total costs ($B + C$), where all inputs and outputs are valued at market prices. A positive private profit indicates the market competitiveness of a commodity system under investigation, given input and output prices, technologies, and government policies.

The second line of the matrix measures social profitability, which is social revenue less social costs. To arrive at the second row of the matrix, social prices (which reflect the underlying scarcity and thus the optimal allocation of resources) are used to value inputs and outputs. A positive social profit indicates that the system uses scarce resources efficiently and contributes to national income (i.e., the commodity has a static comparative advantage).

The second accounting identity, found in the third row of the table, is the difference between the first and second rows. If market-failure correction policies by the government do not exist (or are negligible), any differences between the first row and the second row must be caused by distorting policies. The third row also reflects transfers between producers on one side and government treasury and consumers on the other side. Under certain restrictive assumptions, transfers to producers are identical to producer surplus and transfers to consumers are identical to consumer surplus.⁴

³Unlike sugar, producers receive less than the world market price for molasses. First claim on molasses is by those mills with their own alcohol factories. The second is for use in domestic industry. The remainder is exported through a state-owned marketing board.

Table 2. Price Comparison between Adjusted CIF Price and Wholesale Price for Sugar (1981–86)

Year	Adjusted CIF Price (000 RP/mt)	Wholesale Price (000 Rp/mt)	NPC ^a
1981	349	502	44
1982	207	523	53
1983	296	532	80
1984	219	565	58
1985	195	600	208
1986	347	611	76

Source: Rosegrant et al., p. 4.34, 1987.

Note: The adjusted CIF is the actual CIF price at the port adjusted for internal marketing costs necessary to transport the sugar to the wholesale market.

^aNPC is the nominal protection coefficient.

⁴The assumptions necessary for this to be true include supply and demand elasticities equal to zero and no changes in input prices.

Table 3. An Outline of the Policy Analysis Matrix

	Revenues	Costs		Profits
		Tradable Inputs	Domestic Factors	
Valued at private prices	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i> ^a
Valued at social prices	<i>E</i>	<i>F</i>	<i>G</i>	<i>H</i> ^b
Divergences	<i>I</i> ^c	<i>J</i> ^d	<i>K</i> ^e	<i>L</i> ^f

Source: Based on Monke and Pearson.

^a Private profit, $D = (A - B - C)$.

^b Social Profit, $H = (E - F - G)$.

^c Output Transfers, $I = (A - E)$.

^d Input Transfers, $J = (B - F)$.

^e Factor Transfers, $K = (C - G)$.

^f Net Policy Transfers, $L = (D - H) = (I - J - K)$.

Data and Modeling Assumptions

The sugar production, processing, and marketing system can be represented by three separate activities. The first activity, farm production, consists of cane production and transportation to the mills. Processing is the second activity, with cane as the input and sugar as the output. Transportation of sugar to the wholesale market, at which level the domestic sugar price can be compared with the CIF import price, is the third activity.

The possibility of multiple harvests from a single planting complicates the analysis. Cane usually needs sixteen months from land preparation to harvest. Once cane is harvested, however, roots intentionally left in the ground grow again. This second or "ratoon" crop needs only twelve months to mature. This process can go on for several generations, producing less yield with each subsequent generation. On irrigated area, cane is rarely grown for more than one ratoon, while on dryland area two ratoons are common. Each generation utilizes different amounts of inputs and produces different amounts of output. The farm-level activity thus is modeled as five separate representative farms: S I, S II, T I, T II, and T III, where S I denotes first generation cane planted on irrigated land, T I denotes first generation cane planted on dry-land area, and so on.

Because the Indonesian economy is relatively undistorted, determination of social prices was straightforward. To determine the shadow exchange rate, it is necessary to compensate for overvaluation from two sources, differential inflation and the structure of protection. A devaluation in 1986 eliminated the effects of differential inflation on the exchange rate, and the shadow exchange rate reflects only the distor-

tion associated with the structure of protection, approximately 8% (see Panggabean, p. 28, for details). The Indonesian rupiah is freely convertible, and capital mobility is essentially unlimited. Therefore, the market interest rate for loans to similar agricultural activities was used as the social interest rate (20%). Pearson et al. report that agricultural labor markets function in a relatively competitive manner. Therefore, the market and social wage rate are assumed identical. The private (social) opportunity cost of land is calculated as the difference between the private (social) revenue and variable costs of the crops replaced by cane. The S I and T I cane crops are assumed to take sixteen months, while the S II, T II, and T III crops require only twelve months. The crop combinations used to calculate the opportunity cost of land are (see Monke and Pearson, pp. 210-11 for an example of this type of calculation):

S I—2 irrigated rice, 1 maize, 1 soybean
(first crop, irrigated cane)

S II—2 irrigated rice, 1 maize
(second crop, irrigated cane)

T I—2 dryland rice, 1 soybean
(first crop, dryland cane)

T II/III—2 dryland rice each
(second and third crop, dryland cane).

Data for 1987 were used because they are reasonably representative of expected long-run levels of domestic and world prices of inputs and outputs. Finally, because off-Java area is relatively small and data on production techniques and costs off-Java are not available, the analysis is limited to Java cane production. A technical appendix with further details is available from the authors.

Results and Discussion

The basic results from the farm production activity's budget can be divided into two main categories: cane produced by farmers joining the TRI program (table 4) (75% of the total cane area in Java), and other cane including cane grown by the mills (details not reported here). The only difference between the two is the subsidized interest rate TRI farmers receive.

The results indicate that in 1986/1987 both dryland and irrigated cane cultivation were socially unprofitable.⁵ The principal cause of that result is the opportunity cost of the land. For most of the cultivation techniques, the opportunity cost of land was equivalent to a large share of the value of the sugar produced. It is interesting to note that the social opportunity cost of land is somewhat lower than its private opportunity cost. This result occurs because part of the private return to land in cultivation of alternative crops arises from input subsidies for those crops.

For most sugar growers, sugar cultivation was also privately unprofitable. The only exceptions were the first and second ratoon crops in dryland areas (T II and T III). Once the investment in planting cane in dryland areas is sunk, the private return to ratoon crops is positive. Private profitability is less negative than social profitability as a result of transfers to producers. On the inputs side, the policy of maintaining low prices of chemical inputs, especially fertilizer, to encourage food crop production has been viewed as a major incentive policy for agricultural growth in Indonesia.⁶ For example, the fertilizer subsidy was approximately 30% for urea and 54% for TSP in 1986. Cane producers benefitted from these subsidies (table 5). The major type of fertilizer used in cane production is ammonium sulfate (AS). In 1986, the farm-gate price was 125 rupiah (Rp) per kilogram, while the world price was Rp 196 per kilogram, which implies a subsidy of Rp 49,000 per hectare for S I farmers and slightly smaller amounts for other technologies.

⁵Social profitability accrues to the entire system, not to an element of it. However, to make the social/private cost comparison at the farm level, the social revenue of production is arbitrarily taken to be the farm-level equivalent of the world price times farmers' share of sugar plus the world price of molasses times farmers' share of molasses.

⁶Some observers of Indonesian agriculture have argued that at least a part of the fertilizer subsidy was necessary to offset a market failure in the provision of information about the profitability of the new rice technologies. For the analysis in this paper, all of the fertilizer subsidy is considered a market distortion.

The biggest government direct subsidy on inputs came in the form of the credit subsidy. TRI farmers have access to 12% credit, while the social interest rate is 20% per annum. This differential provides a Rp 146,000 per hectare subsidy for the S I crop, with the T III crop as the lowest with Rp 60,000 per hectare. For all technologies, the credit subsidy alone is substantially larger than the subsidy on fertilizers. Together, the credit and fertilizer subsidies are the prime determinants of the transfer on the inputs side.

The principal determinant of transfers to the farm production activity is the difference between world and domestic prices. In 1986/1987, farmers received a provenue price of Rp 467.50 per kilogram for their share of sugar produced (approximately 62%).⁷ The CIF price for sugar in 1987 was Rp 329 per kilogram (approximately 1C¢/lb), equivalent to a mill-gate social price of Rp 329 per kilogram (after converting to social costs and subtracting the social value of transport costs). For one hectare of S I cane, private gross revenue was Rp 2.73 million, while social gross revenue was only Rp 2.01 million.

The aggregate effects of government policies on sugar production can be calculated by multiplying the per hectare results by the number of hectares in each type of cane production (table 6). In 1986/87, the foregone social net revenue was Rp 248 billion. Farmers bore part of this cost in the form of a private loss of Rp 115 billion. The remainder was borne by consumers and the government in the form of higher consumer prices and government subsidies on various inputs.

To gain further insight into the effect of the opportunity cost of land, it is useful to decompose the aggregate PAM into two matrices, one each for cane cultivation on irrigated and on dryland areas (table 7). Social profitability in the irrigated areas is negative primarily because of the high social opportunity cost of land. For dryland areas, social profitability is still negative, but private profitability is positive. Even though the first crop in the dryland areas has negative profit, the positive profitability of the ratoon crops more than compensates. This result is supported by anecdotal evidence that sugar production in dryland areas is privately profitable and that it is quite common in those areas to find sugar that has been ratooned more than once.

⁷Two percent of farmers' share is received in kind and valued at the ex-factory price.

Table 4. The Policy Analysis Matrix for TRI Sugar Cane Production in Java, 1986/1987 (000 rupiah/hectare)

Item	S I ^a	S II	T I	T II	T III
Private					
Labor	615	418	525	377	369
Capital	220	108	171	91	89
Tradables	444	250	373	212	206
Land and management	2,547	1,993	1,312	758	758
Total cost	3,826	2,769	2,381	1,439	1,423
Revenue ^b	2,731	2,211	1,877	1,692	1,542
Profit	-1,094	-558	-504	253	119
Social					
Labor	615	418	525	377	369
Capital	366	180	285	152	149
Tradables	560	303	460	259	253
Land and management	2,161	1,885	985	709	709
Total cost	3,702	2,786	2,256	1,497	1,480
Revenue ^c	2,009	1,630	1,382	1,246	1,137
Profit	-1,693	-1,157	-873	-251	-344
Policy effects (private minus social)					
Total cost	123	-17	125	-59	-57
Labor	0	0	0	0	0
Capital	-146	-72	-114	-61	-60
Tradables	-116	-53	-88	-47	-47
Land and management	386	108	327	49	49
Revenue	722	581	495	445	405
Net effects	599	598	370	504	463

Source: Own estimates; data appendix available from authors.

Note: Some numbers do not add up because of rounding errors.

^a S I—planted cane, grown on irrigated areas; S II—first ratoon cane, grown on irrigated areas; T I—planted cane, grown on dry land areas; T II—first ratoon cane, grown on dry land areas; and T III—second ratoon cane, grown on dry land areas.

^b Private revenue is the provenue price times 98% of farmers' share of sugar, plus ex-factory price times 2% of farmers' share, plus domestic price of molasses times farmers' share of molasses.

^c Social revenue accrues to the whole production and processing chain. For this table, social revenue of production is arbitrarily taken to be the farm-level equivalent of the world price times farmers' share of sugar, plus world price of molasses times farmers' share of molasses.

Processing

The total private cost of milling cane (excluding the cost for the cane) was Rp 383 billion (table 8), while total revenue was only Rp 372 billion, resulting in a private loss of Rp 11 billion. The social cost of sugar processing was substantially larger than the private cost. The principal source of this difference was the capital subsidy to the

mills of Rp 69 billion. The private cost of capital to the mills was assumed to be only 10% (because they had access to foreign, low cost loans for rehabilitation) while the social cost was 20%. Social revenue came mainly from the mills' share of sugar output (37.2%). Transfer to the mills on the revenue side was Rp 75 billion. The net transfer to the mills was Rp 159 billion.

Post-Factory Transportation

The final category of cost is for transportation from factories to the wholesale level. The average cost of transporting a kilogram of sugar was Rp 25 (Rosegrant et al., p. 4.42); the total private cost of transportation from mills to the wholesale market is Rp 44 billion. The Indonesian input-output table for 1983 was used to break down the costs into the various categories (table 9).

Table 5. Major Input Subsidies for TRI farmers, 1986/1987 (000 rupiah/hectare)

Type	S-I	S-II	T-I	T-II	T-III
AS	49	49	44	43	43
TSP	41	0	29	0	0
KCL	12	0	0	0	0
Capital	146	72	114	61	60

Source: Own estimates; data appendix available from authors.

Table 6. Total Transfer in the Farm Activity, 1986/1987 (billion rupiah)

	Labor	Land and Mngmt.	Capital	Tradable Inputs	Total Cost	Revenue	Profit
Private	120	421	41	77	660	545	-115
Social	120	373	60	96	649	401	-248
Divergences	0	48	-19	-19	10	144	133

Source: Own estimates; data appendix available from authors.

Table 7. Policy Analysis Matrix by Land Type, 1986/1987 (billion rupiah)

	Labor	Land and Mngmt.	Capital	Tradable Inputs	Total Cost	Revenue	Profit
Irrigated							
Private	72	318	24	49	462	346	-117
Social	72	281	41	62	456	255	-201
Divergences	0	35	-15	-12	8	91	82
Dryland							
Private	48	104	17	29	198	199	1
Social	48	92	22	36	198	147	-51
Divergences	0	13	-4	-7	2	52	50

Source: Own estimates; data appendix available from authors.

Table 8. The Policy Analysis Matrix for the Mills in Java, 1987 (billion rupiah)

	Labor	Land	Capital	Tradable Inputs	Total Cost	Revenue	Profit
Private	121	0	69	193	383	372	-11
Social	121	0	138	208	467	297 ^a	-170
Divergences	0	0	-69	-15	-84	75	159

Source: Own estimates, CBS; data appendix available from authors.

Note: Numbers might not add up because of rounding errors.

^a Arbitrary allocation of a share of the social revenue of the whole system to the processing sector. See also the note to table 4.**Table 9. Total Transfer in the Transportation Activity, 1986/1987 (billion rupiah)**

	Labor	Land	Capital	Tradables Cost	Total	Revenue	Profit
Private	13		12	15	44	44	0
Social	13		12	21	46	46	0
Divergences	0		0	-1	-1	-1	0

Source: Own estimates; data appendix available from authors.

The Indonesian Sugar Policy Analysis Matrix

The net effects of policy in the sugar system are shown in the PAM in table 10. In terms of both private and social profitability, sugar production in Indonesia is unprofitable. (See Pearson et al.

for a similar result.)⁸ The social loss is Rp 465 billion, about 1% of agricultural value added. Farmers, consumers, and the government all

⁸Strictly speaking these results are for Javanese production only. See earlier comments about the role of production off-Java.

Table 10. Policy Analysis Matrix for the Sugar Commodity System in Indonesia, 1987 (billion rupiah)

	Labor	Land	Capital	Tradables	Total Cost	Revenue	Profit
Private	255	421	122	290	1,088	961	-127
Social	255	373	210	324	1,162	698	-465
Divergences	0	48	-88	-34	-75	263	338

Source: Tables 6 and 8.

subsidize the production of sugar. For farmers, the transfers received from subsidized capital, inputs, and higher-than-world sugar prices are offset by the losses from not being able to grow more profitable alternative crops. Consumers pay Rp 263 billion more than they would if the domestic sugar price were equal to the world price. The government loses Rp 88 billion in the form of a capital subsidy to the mills and TRI farmers, and Rp 34 billion for subsidies on chemical inputs.

Conclusions

Food policy analysts normally think in terms of trade-offs between consumers and producers, but Indonesian sugar policy does not fit that model. Instead, producers, consumers, and the government budget all have paid to help the government advance its objective of self-sufficiency in sugar. The policy analysis matrix is an effective tool to sort out the conflicting effects arising from the various policy instruments used to achieve that objective. Other measures such as the DRC or the EPC would illuminate only part of the picture.⁹ For example, the DRC would reveal the negative social profitability of cane production but would not make clear the complex set of transfers that arise. The EPC would indicate the impact of product market transfers but not show factor market transfers or measure social profitability.

Policy makers know what policy goals are and what policy instruments are being used. The PAM provides a framework in which to measure the costs and benefits of those policy instruments to various market participants and to the economy, and the results are presented quickly and efficiently. At the same time, the limitations of the PAM should be made clear. The PAM typically

uses fixed input-output coefficients, so it is not possible to use the PAM directly to indicate producer or consumer responses to policy changes that reduce distortions. Transfers to consumers and producers measured by the PAM are exactly equal to the more familiar consumer and producer surplus measures only under restrictive assumptions. However, the transfers are roughly equal to the rectangles associated with traditional surplus calculations.

The instruments used to implement Indonesian sugar policy are complicated and contradictory, with some partially offsetting the effects of others. The net effects are unusual, even by the standards of sugar policy around the world. Domestic output prices are above world prices and domestic input prices are lower than world prices, a situation that normally results in excess profits for cane producers. In spite of these price policies, Javanese farmers must be compelled to grow cane. Because the profitability of alternate crops is so much higher than cane, even substantial government and consumer transfers to producers were not enough to induce farmers to grow cane willingly.

Until 1989, Indonesia was essentially self-sufficient in sugar, but this basic policy goal was met at substantial cost to producers, consumers, and the government. Farmers, forced to grow cane on irrigated land, had negative private profits, and the Indonesian economy lost because the social cost of cane production (and milling) was much higher than its social value. A large share of cane production is on irrigated land that otherwise would be used to grow crops, primarily rice, with much higher profit both privately and socially. By mandating cane production, the potential for socially profitable rice self-sufficiency (the most important Indonesian agriculture policy goal) is reduced. The government also lost financially because of the credit and fertilizer subsidies, which reduce the losses incurred by farmers and the milling firms. Consumers lost because domestic prices were set higher than world prices, again to reduce the

⁹These measures can be calculated easily from the values in the PAM. See Monke and Pearson.

losses incurred by farmers. Only a few cane growers using dry land benefitted.

Indonesia's sugar policies reflect the history of sugar in that country. Succeeding generations of policy makers have made incremental changes to the existing policy set rather than start anew. For example, the most recent change—the TRI program, which ended land rental by the mills and mandated that production be managed by smallholders on their own land—left in place many of the policies of the past. As a result, a complex web of policies has grown up with a correspondingly complex underpinning of interest groups (e.g., employees and owners of socially unprofitable sugar mills, sugar cane dealers and farm managers, sugar traders) that benefit from various aspects of the policies. Any change either must balance the interests of the current beneficiaries with those of the much more numerous sugar consumers and producers who would gain from reform, or else result from a powerful, exogenous shock.

[Received April 1990; final revision received October 1990.]

References

- Anonymous (Tim Survei). "Bahan Seminar Tim Survei dan Analisa Produksi dan Konsumsi Gula dan Pemanis Lain." Jakarta, Indonesia: Dewan Gula Indonesia, 1987.
- Biro Pusat Statistik (BPS). *Economic Census of Large and Medium Manufacturing Establishment Sector*. Jakarta, Indonesia, 1987.
- . *Penyusunan Tabel Input Output Indonesia Updated 1983*. Jakarta, Indonesia, 1985.
- . *Struktur Ongkos Usaha Tani Padi dan Palawija 1986*. Jakarta, Indonesia, 1988.
- Balassa, Bela, and D. M. Schydrowsky. "Domestic Resource Costs and Effective Protection Once Again." *J. Polit. Econ.* 80(1972):63–69.
- Blackburn, Frank G. *Sugar Cane*. London: Longman, 1984.
- Brown, Colin. "Specialisasi Regional dalam Menghasilkan Komoditi: Pengalihan Industri Gula ke Luar Jawa." *Prisma* 17(1988):116–28.
- Bruno, Michael. "Domestic Resource Costs and Effective Protection: Clarification and Synthesis." *J. Polit. Econ.* 80(1972):16–33.
- . "The Optimal Selection of Export-Promoting and Import-Substituting Projects." *Planning the External Sector: Techniques, Problems and Policies, Report on the First Interregional Seminar on Development Planning*. New York: United Nations, 1967.
- Corden, W. Max. *The Theory of Protection*. London: Oxford University Press, 1971.
- Falcon, Walter P., William O. Jones, Scott R. Pearson, John A. Dixon, Gerald C. Nelson, Frederick C. Roche, and Latrian J. Unnevehr. *The Cassava Economy of Java*. Stanford CA: Stanford University Press, 1984.
- Gittinger, J. Price. *Economic Analysis of Agricultural Projects*, 2nd ed. Baltimore MD: Johns Hopkins University Press, 1982.
- Husz, Georg S. *Sugarcane: Cultivation and Fertilization*. Bochum, West Germany: Ruhr-Stickstoff A.G., 1972.
- Krueger, Anne O. "The Political Economy of the Rent-Seeking Society." *Amer. Econ. Rev.* 5(1974):291–303.
- Monke, Eric A., and Scott R. Pearson. *The Policy Analysis Matrix for Agricultural Development*. Ithaca NY: Cornell University Press, 1989.
- Mubyarto. "Tebu Rakyat Intensifikasi: Prospek dan Masalahnya." *Prisma*, no. 10 (1981), pp. 50–61.
- . "The Sugar Industry." *Bull. Indonesian Econ. Stud.*, no. 2(1969), pp. 37–59.
- . "The Sugar Industry: From Estate to Smallholder Cane Production." *Bull. Indonesian Econ. Stud.*, no. 2 (1977), pp. 29–44.
- Panggabean, Martin P. *Measuring the Cost of Indonesian Sugar Policy: The Policy Analysis Matrix Approach*. M.S. thesis, University of Illinois, 1989.
- Pearson, Scott R., Narongchai Akrasanee, and Gerald C. Nelson. "Comparative Advantage in Rice Production: A Methodological Introduction." *Food Res. Inst. Stud.* 15(1976):127–37.
- Pearson, Scott R., Walter P. Falcon, Paul Heytens, Eric Monke, and Roz Naylor. *Rice Policy in Indonesia*. Ithaca NY: Cornell University Press, forthcoming.
- Rosegrant, Mark W., Faisal Kasryno, Leonardo A. Gonzales, Chairil Rasahan, and Yusuf Saefudin. *Price and Investment Policies in the Indonesian Food Crop Sector*. Washington DC: International Food Policy Research Institute, 1987.
- Sapuan, et al. *Ekonomi Pergulaan di Indonesia*. Jakarta, Indonesia: Badan Urusan Logistik, 1985.
- Sievers, Allen M. *The Mystical World of Indonesia, Culture and Economic Development in Conflict*. Baltimore MD: Johns Hopkins University Press, 1974.
- Soetrisno, Noer. "Farmers, Millers and Sugar Production in Indonesia." Ph.D. thesis, University of the Philippines Los Banos, 1984.

Education and Innovation Adoption in Agriculture: Evidence from Hybrid Rice in China

Justin Yifu Lin

This paper uses the diffusion of F_1 hybrid rice as a case for examining the effects of education on the adoption of new technology in China. A simple behavioral model that treats the adoption of hybrid rice as a portfolio selection problem is presented. The implications of the model are tested with farm-level data collected from a sample of 500 households in Hunan Province. The results from a dichotomous probit model and a two-limit tobit model are consistent with the hypothesis that education has a positive impact on the adoption of new technology.

Key words: China, economic development, education, portfolio-selection model, technology diffusion.

The continuous creation and introduction of new technology has been used as a standard for distinguishing a modern agricultural system from a traditional one (Schultz 1964). However, the introduction of many new technologies has met with only partial success, as measured by observed rates of adoption. Constraints to the rapid diffusion of a new technology may arise from many sources, such as lack of credit, inadequate farm size, unstable supply of complementary inputs, and so on (Feder, Just, and Zilberman). This paper examines the role of education in a farm household's decision regarding adoption of F_1 hybrid rice seed in the context of the Chinese economy.

A new agricultural technology may reflect high yield, low cost, or other desirable traits. However, the changes in the production process involved in the adoption of a new technology may

bring risks resulting from imperfect information and the possibility of committing errors. Because education enhances one's ability to receive, decode, and understand information, Schultz (1964, 1975) and Nelson and Phelps, among others, have hypothesized that education may facilitate the diffusion of new technology. That is, farmers with a relatively high level of education may have a higher probability of adopting new technologies than those with relatively little education.

A number of empirical studies have examined the linkage between adoption of new agricultural technology and education. Some of the evidence is summarized by Feder, Just, and Zilberman. In those studies, the likelihood of adopting a new technology was found to be positively related to the education level of farmers. A recent study by Duraisamy also found that in India the level of using high-yielding rice varieties is positively related to education level. The above theory and evidence suggest that a farmer's education level may be an important factor in the adoption of F_1 hybrid rice in China.

China is the only country in the world in which F_1 hybrids are commercially used in production. Several studies have estimated the profitability of hybrids as compared to conventional varieties (He et al. 1984; He, Zhu, and Flinn 1987a, b). The yield advantage of hybrids over conventional varieties is about 15%. Because China's economy is centrally planned, governmental intervention has often been used in promoting cer-

Justin Yifu Lin is an associate professor of economics at Peking University and a visiting associate professor of economics at the University of California at Los Angeles. He is also an adjunct professor at Australian National University.

This study was supported by the Rockefeller Foundation, Grant No. 880-0489.

The paper was prepared for the Final Workshop on Differential Impact of Modern Rice Technology on Favorable and Unfavorable Production Environments held 26–28 March 1990 at the International Rice Research Institute.

The author gratefully acknowledges valuable comments by Yujiro Hayami, Peter Warr, Keijiro Otsuka, Cristina David, Mark Rosegrant, and other participants of the workshop. Shen Minggao and Wang Fang provided very capable assistance in data collection and processing. Robert Ashmore gave a helpful exposition review. The author is also indebted to two anonymous referees for helpful comments on an earlier draft of this paper.

tain technologies. As a result, just how important economic considerations are in the adoption decision regarding F_1 hybrids at the farm level is a controversial question (Wiens; Barker and Herdt, p. 61). An earlier study of my own (Lin 1991), using aggregate panel data, found profitability to be an important factor in explaining differences across regions and over time in the diffusion of hybrids after the household-based farming system reform, but not in the prereform period.¹ However, because of the aggregate nature of the data set used in that study, direct conclusions about the role played by education and other household characteristics could not be drawn.

For the purpose of this study, cross-section production data collected at the end of 1988 from a sample of 500 households are used. The main purpose is to determine the role of education in a household's decision regarding adoption and intensity of use of F_1 hybrids. In addition, other potentially relevant variables, such as the farming experience and sex of the household head and the availability of credit, are included.

The paper is organized as follows: The next section presents a simple behavioral model, which treats the adoption decision as a portfolio selection problem. The model is followed by a description of the areas covered by the study and the data set. The subsequent section presents the empirical analysis. The last section summarizes the results and discusses their implications.

A Portfolio Selection Model of Technological Adoption

In order to analyze the impact of education on the adoption decision regarding F_1 hybrid seeds, this section constructs a formal behavioral model in which the adoption of new technology, represented by the use of F_1 hybrid seed, is treated as a portfolio selection problem. The objective of a farm household is to maximize its expected utility. Following the convention of portfolio selection literature, a household is assumed to care only about the mean and variance of its income.² For simplicity, a farm household grows only a single crop, rice, on a unit of land. There

are two alternative technologies—conventional rice (CR) and F_1 hybrid rice (HR). CR is assumed to have a low yield, but the output is certain to each household. The output level, however, may be different from household to household as a result of the difference in some household characteristics. HR, on the other hand, has a higher expected yield but the output level is uncertain to each household. The variance of the output level will be assumed to depend on the i th household head's education level and other household-specific variables that affect the household's ability in dealing with new technology.³

Following Jamison and Lau's convention (p. 197), if CR is used on all of the i th household's land, its income can be written as a function of variables representing the economic environment and a specific household factor, in the form

$$(1) \quad Y_{i,C} = M_C(E) + \epsilon_{i,C},$$

where E is a vector of independent variables representing the prices of rice, seeds, chemical fertilizers, pesticides, and so on, and $\epsilon_{i,C}$ is a variable representing the i th household's specific capacity for producing CR.⁴ Similarly, the mean income from using HR on all its land can be written as a function of the same independent variables, and a specific household factor, in the form

$$(2) \quad Y_{i,H} = M_H(E) + \epsilon_{i,H}.$$

The variables $\epsilon_{i,C}$ and $\epsilon_{i,H}$ cannot be directly observed, but it is assumed that their joint distribution over the whole population can be described by a probability density function. Therefore, the i th household's mean income with r_i proportion of its land producing HR can be expressed as

$$(3) \quad Y_i = \{M_C(E) + \epsilon_{i,C} + r_i[D(E) + (\epsilon_{i,H} - \epsilon_{i,C})]\},$$

where $D(E) \equiv M_H(E) - M_C(E)$.

If a household allocates all its land to produce CR, there will be no variance in its income, as the output level of CR is assumed to be constant to each household. If a household allocates all

¹ Farming activities were organized in a collective team system prior to the farming institutional reform in 1979. The collective system was replaced by an individual household-based system by 1983. For a theoretic and empirical study of this reform, see Lin (1988).

² This assumption requires that the utility function be quadratic, or/and that the outputs of new as well as old technologies have a normal distribution.

³ These assumptions imply that the adoption of CR is riskless, and that HR is risky.

⁴ That is, the first component, $M_C(E)$, is common to each household, and the second component, $\epsilon_{i,C}$, differs from household to household. The difference in the second component may arise from several sources, like the quality of irrigation, fertility, and other microphysical conditions of the household's landholding, the experience and ability of the household members, and so on.

its land to produce HR, the variance in the household's income will be assumed to have the following functional form:

$$(4) \quad V_i = V_{i,H}(e_i, Z_i, E),$$

where e_i is the i th household head's education level; and Z_i is a vector of other household-specific variables, such as the availability of credit and extension service to the household.⁵ If a household allocates r_i proportion of its land to produce HR, the variance in its income will thus be

$$(5) \quad V_i = r_i^2 \cdot V_{i,H}(e_i, Z_i, E).$$

From expression (5), we find that the variance of a household's income is positively related, and increases at an increasing rate, to the proportion of land that the household allocated to HR.

For simplicity and ease of interpretation, the utility function is assumed to be separable, and to have the following specific form:

$$(6) \quad U_i = Y_i - C(V_i) = Y_i - C(r_i, e_i, Z_i, E),$$

where $C_1 > 0$, $C_{11} > 0$, $C_{12} < 0$, and $C(0, e_i, Z_i, E) = 0$.

The specifications of $C(\cdot)$ in the second line of expression (6) imply that, given other variables, the utility loss increases and at an increasing rate with the ratio of land allocated to HR, that education reduces the utility loss of adopting HR, and that there is no utility loss if HR is not adopted. These characteristics can be justified by the previous assumptions that the income variance increases and at an increasing rate with the proportion of land used in HR, education reduces the variance of HR output, and

ity in producing CR and HR, and other exogenous variables in the household as well as economy, the household will allocate its land to HR in a way that its utility is maximized.

The Probability of Adoption

Before solving for the optimal proportion of land used for HR, r_i^* , we will first consider the probability that the i th household will allocate part of its land endowment to HR, that is, r_i will be greater than zero. Expression (7) indicates that the necessary condition for $r_i > 0$ is $r_i \cdot [D(E) + \epsilon_{i,H} - \epsilon_{i,C}] - C(r_i|e_i, Z_i, E) > 0$. Therefore, whether a household adopts HR depends only on the value of $r_i \cdot [D(E) + \epsilon_{i,H} - \epsilon_{i,C}] - C(r_i|e_i, Z_i, E)$. As shown in figure 1, the necessary and sufficient condition for $r_i > 0$ is

$$(8) \quad C_1(0|e_i, Z_i, E) < D(E) + \epsilon_{i,H} - \epsilon_{i,C},$$

where $C_1(0|e_i, Z_i, E)$ is the first derivative of $C(r_i|e_i, Z_i, E)$ evaluated at $r_i = 0$. The adoption indicator of the i th household, A_i , takes the value of

$$(9) \quad A_i = \begin{cases} 1 & \text{if } C_1(0|e_i, Z_i, E) < D(E) + \epsilon_{i,H} - \epsilon_{i,C}, \text{ HR is adopted, and} \\ 0 & \text{if } C_1(0|e_i, Z_i, E) \geq D(E) + \epsilon_{i,H} - \epsilon_{i,C}, \text{ HR is not adopted.} \end{cases}$$

Thus, the probability that a household drawn randomly from the population, with education of household head, other characteristics and environment variables given, would adopt HR equals

$$\begin{aligned} P_i &= \Pr(A_i = 1) = \Pr[C_1(0|e_i, Z_i, E) < D(E) + \epsilon_{i,H} - \epsilon_{i,C}] \\ &= \Pr[C_1(0|e_i, Z_i, E) - D(E) < \epsilon_{i,H} - \epsilon_{i,C}]. \end{aligned}$$

the output of CR is constant. The optimization problem for a household can thus be expressed as follows:

$$(7) \quad \begin{aligned} \text{Max}_{0 \leq r_i \leq 1} \quad & U_i(r_i|e_i, \epsilon_{i,C}, \epsilon_{i,H}, Z_i, E) \\ & = \{M_C(E) + \epsilon_{i,C} + r_i \cdot [D(E) + (\epsilon_{i,H} - \epsilon_{i,C})]\} - C(r_i|e_i, Z_i, E). \end{aligned}$$

This expression postulates that, given the i th household head's education level, specific abil-

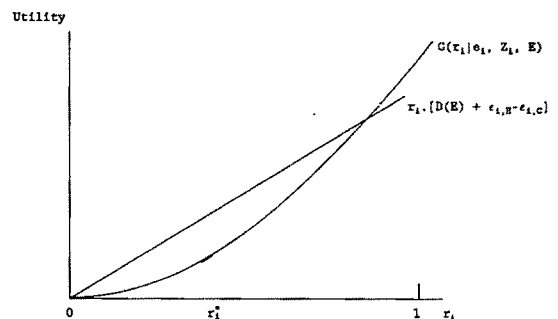


Figure 1. The optimal land allocation to hybrid rice

⁵ Because education enhances the ability of using new technology, the variance of HR output will be postulated to be inversely related to the educational level of the household head.

This probability depends on the difference of the functions $C_1(0|e_i, Z_i, E)$ and $D(E)$ and $\epsilon_{i,H} - \epsilon_{i,C}$. $C_1(0|e_i, Z_i, E)$ is specified as a linear function of e_i ; that is,

$$C_1(0|e_i, Z_i, E) = a_0 e_i + a_1 Z_i + a_2 E,$$

where a_0 is an unknown parameter, and a_1 and a_2 are both row vectors of unknown parameters. Similarly, $D(E)$ is a linear function of E ,

$$D(E) = bE,$$

where b is a row vector of unknown parameters. We will denote the difference in the individual household factors as ϵ_i ; that is,

$$\epsilon_{i,H} - \epsilon_{i,C} \equiv \epsilon_i.$$

Then, the probability of the i th household adopting HR equals

$$(10) \quad P_i = \Pr(A_i = 1) = \Pr[\epsilon_i < bE - a_0 e_i - a_1 Z_i - a_2 E] \\ = F(b'E - a_0 e_i - a_1 Z_i),$$

where $b' = b - a_2$, and $F(\cdot)$ is the cumulative distribution function.

As shown, the probability of the i th household adopting HR seed is the value of the cumulative distribution function of F evaluated at $b'E - a_0 e_i - a_1 Z_i$. The exact distribution of F depends on the population distribution of the random variable ϵ_i . If ϵ_i is identically and independently distributed as the normal distribution over the population, the unknown parameters b' , a_0 , and a_1 can be estimated by the probit regression model, which yields consistent and asymptotically efficient estimates. Since $C_{12} < 0$ implies that $a_0 < 0$, a larger e_i thus implies a higher probability of adopting HR seed.

Optimal Adoption Intensity

The i th household's optimal proportion of land used in the production of HR can be obtained by solving expression (7). The first-order condition for an optimum requires

$$(11) \quad D(E) + \epsilon_{i,H} - \epsilon_{i,C} - C_1(r_i|e_i, Z_i, E) = 0.$$

Equation (11) implies that, for the optimality to hold, the decision maker equates the utility gain from adopting HR with the utility loss resulting from adopting this new technology at the margin. The second-order condition requires that $-C_{11} < 0$, which holds according to the basic assumptions.

The effect of education on the optimal proportion of land used for HR is implicitly defined in equation (11). From the implicit function theorem, we can obtain the following relation:

$$(12) \quad dr_i/de_i = -C_{12}/C_{11} > 0.$$

The optimal proportion of a household's land used for HR is thus implied in expression (11). To be specific, the optimal proportion can be expressed in a functional form as follows:

$$r_i^* = r(e_i|\epsilon_i, Z_i, E),$$

where $\epsilon_i = \epsilon_{i,H} - \epsilon_{i,C}$, as previously defined, is assumed to be identically and independently distributed as the normal distribution over the population. We will assume that the function is linear in e_i , Z_i , and E , that is,

$$(13) \quad r_i^* = c_0 e_i + c Z_i + dE + \epsilon_i,$$

where c_0 is an unknown parameter, and c and d are row vectors of unknown parameters. Because $0 \leq r_i \leq 1$, the dependent variable r_i is censored. The tobit regression model with lower bound zero and upper bound 1 will yield consistent and asymptotically efficient estimates of the unknown parameters in expression (13).

The discussion so far focuses solely on the effect of a household head's education on the probability and intensity of adopting HR. It is also possible to investigate the impacts of other household characteristics and environmental variables under the present framework. For example, an increase in the price of rice will increase the relative profit of HR over CR and, therefore, increase the probability and intensity of adopting HR. Changes in the price of purchased inputs may also affect the probability and intensity of adopting HR if the requirements of purchased inputs differ between CR and HR. The availability of credit may facilitate adoption if more purchased inputs are used for HR and a household's liquidity becomes a constraint. Furthermore, most farm households are obliged to sell a certain quota of rice at below-market price to the government. In cases where the quota is a binding constraint, a household may adopt hybrid rice simply for its higher yield, even though it may not be as profitable as conventional rice. However, a full-length discussion of the impacts

of these other variables is beyond the purview of this paper.⁶

The Data and Study Setting

The data come from a cross-section survey of 500 households in five counties in Hunan Province which was carried out during December 1988 and January 1989. Hunan Province is located on the middle reaches of the Yangtze River in South China. It has a semitropical climate. The average temperature is 4–7 °C in January and 26–30 °C in July, with 260–300 frost-free days. Annual rainfall is about 1,300 millimeters to 1,700 millimeters. Most of the rainfall is concentrated in May, June, and July. Of the total land area of 211,000 square kilometers, mountains make up 51.2%, hills 29.2%, plains 13.1%, and water surface 6.4%. The province has 2.56 million hectares of cultivated land, 82% of which is irrigated. The per capita cultivated area is 0.05 hectare, below the national average. The total population is 5.8 million, of which 4.8 million is agricultural. From the proportion of agricultural population in total population, and from the per capita gross values of agricultural output and industrial output, Hunan can be considered in China to be a predominately agricultural province (see table 1). Rice is its most important crop. In 1987, 57% of the total cultivated acreage (or 82% of the grain acreage) was planted with

rice, and, of the rice acreage, 46.5% was planted with hybrid seed.

The province is divided juridically into 105 counties in three types of geographic setting—lake-plain, hill, and mountain. Among the five counties in the data set, the first two—Tiaojiang and Xiangxiang—are selected from the hill region, the next two—Nanxian and Anxiang—from the lake-plain region, and the last one—Zhijiang—from the mountain region. These five counties were selected from the provincial sample of thirty-four counties surveyed annually by the State Investigation Team. Table 1 indicates that the 1988 per capita GNPs of these five counties were all lower than the provincial average, but that their per capita gross values of agricultural output were all higher than the provincial average. However, their rice economies are considered typical, and their agricultural conditions and infrastructure are representative of their respective regions.

Samples of 100 households each from these five counties were surveyed. These households were all included in the random samples surveyed by the State Investigation Team. Table 2 summarizes the key characteristics of the samples in each of the five counties. While households in the two hill counties, Tiaojiang and Xiangxiang, had the highest per capita income in 1988, households in the two lake-plain counties, Nanxian and Anxiang, had the largest farm size. The main reason for the large farm size in the lake-plain region is that a substantial amount of cultivated land has been newly reclaimed from Dongting Lake, one of the five largest lakes in China.

Education in this study refers to a household head's years of formal schooling, including general and vocational training. China's school sys-

⁶ The above model is obviously oversimplified. The framework could be further developed in many directions; however, the simplest form of the model is adopted in order to illustrate clearly its implications. For other models that treat technology adoption as dichotomous choice, see Jamison and Lau; as a portfolio selection problem, see Feder, and Just and Zilberman.

Table 1. Economic Profiles of the Study Areas

	Population		GNP Per Capita (Yuan) ^a	Gross Value Per Capita	
	Total	Agriculture (1,000)		Agriculture (Yuan)	Industry
Nation	1,096,140	552,450	1,278	535	1,662
Hunan	59,157	50,356	987	512	983
Tiaojian	765	699	836	701	494
Xiangxiang	843	761	778	631	851
Nanxian	678	573	842	742	723
Anxiang	522	429	761	660	515
Zhijiang	316	286	713	621	479

Source: *China Statistical Yearbook*, pp. 28, 51, 87, 742, 1989.

Hunan Statistical Yearbook, pp. 375–8, 395–8, 407–10, pp. 435–38, pp. 491–2, 1989.

^a US\$1 = 3.7 Yuan in 1988.

Table 2. Characteristics of Sample Farm Households

	Tiaojian (N = 100)	Xiangxiang (N = 100)	Nanxian (N = 100)	Anxiang (N = 100)	Zhijiang (N = 100)
Mean farm size (mu) ^a	5.0	4.9	8.1	8.4	5.9
Mean household size (person)	4.28	4.26	4.59	4.60	4.20
Per capita income (Yuan) ^b	569	607	430	492	463
Share of income from:					
(i) Nonfarm activities (%)	39	21	19	20	29
(ii) Sideline and animal husbandry (%)	31	25	18	21	23

^a 15 mu = 1 hectare.^b In 1988, US\$1 = 3.7 Yuan.

tem consists of primary, secondary, and tertiary level. Because of a strong urban bias in the school system, students in rural areas are at a disadvantage to compete with students in urban areas. Therefore, the dropout rate in rural primary school is high, and only a small portion of primary school graduates continues higher education.⁷ In the samples, only one household head completed college education and 93.3% of household heads have less than 10 years of schooling. The average years of schooling is 5.52, about the level of a primary school graduate, with a large variation across age cohorts. The average for the cohort of household heads with age less than 30 is 7.85, while for the cohort with age 50 or older the average is 3.91. The average years of schooling is 6.39 for the cohort of thirties and 5.37 for the cohort of forties. Thus the younger a household head is, on average, the better is his/her education.

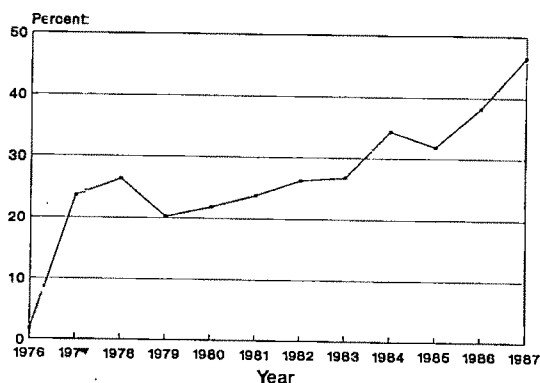
F_1 hybrid seeds were released to farmers in 1976. The price of hybrid rice matched the price of conventional rice. However, the price of hybrid seed was set officially to be ten times the price of conventional seed because initially the seed field produced about one-tenth the seed of a conventional rice field.⁸ Increase in seed costs, however, is mitigated by reduced seed requirements because of low plant density necessary in hybrid rice production.⁹

The yield of F_1 hybrids is reported to be about

20% higher than the conventional rice. In the first two years, the diffusion of hybrid rice in Hunan was very rapid, but it declined sharply in 1979 and stagnated until 1983 (see fig. 2). Several factors may contribute to this fluctuation and stagnation. Substantial government intervention was used at the beginning stage. Because rice-growing environments were diverse but the range of available hybrid varieties initially was limited, crop failures occurred in a number of areas owing to lack of resistance to local diseases. In addition, the growing time of the early-released hybrids was rather long (135 days), and the cooking quality was also a concern. However, most of the aforementioned problems were largely solved by 1988, the time of the survey.¹⁰

The new household-based farming system was introduced in the study areas during 1981–82. Table 3 compares the percentages of households in each county that adopted and did not adopt

¹⁰ The increase in the diffusion of hybrid rice after 1983 can partly be attributed to the improvement in a farmer's incentive for acquiring new innovation resulting from the change from the collective system to the household-based farming system (Lin 1991).

**Figure 2. Diffusion of hybrid rice in Hunan**

⁷ For an informative discussion of the issues and changes in China's rural education, see Perkins and Yusuf, chap. 8.

⁸ The production of F_1 hybrid seed involves a complicated three-line method: (a) locating a cytoplasmic male-sterile parent plant; (b) crossing it with a maintainer line to produce offspring with sterility but with desirable genetic characteristics; and (c) crossing these seeds with a "restorer" line to produce F_1 seeds with normal self-fertilizing power. Initially the yield of hybrid seed fields was very low. The yield has been improved and reached one-third to one-half the yield of a regular field but the price ratio is still maintained.

⁹ Because of high-tillering rate, F_1 hybrids require only one-third to one-fourth of conventional rice's seeding rate.

Table 3. Changes in Adoption of Hybrid Seeds

	Tiaojian (N = 100)	Xiangxiang (N = 100)	Nanxian (N = 100)	Anxiang (N = 100)	Zhijiang (N = 100)
In 1981–82					
Adopter	27	10	4	20	52
Nonadopter	63	90	96	80	48
In 1988					
Adopter	78	67	64	93	99
Nonadopter	22	33	36	7	1

F_1 hybrids at the time when the new farming system was introduced and in 1988. Except in Zhijiang, the majority of households did not adopt hybrids when the institutional change occurred. In contrast, most households in all five counties adopted hybrids in 1988. The higher incidence of early adoption in Zhijiang county probably reflects the fact that most households in Zhijiang grow only one crop of rice a year, while in the other four counties most households grow two crops of rice annually. Therefore, the early-released hybrids, with a rather long maturation period, could be integrated more easily into the cropping system in Zhijiang.

Among the 500 households in the sample, 78% reported to have increased their hybrid-planted acreage, while only 4.2% reported that their hybrid acreage was reduced. The main reason reported for increases in hybrid rice acreage was improvement in yield advantage (384 out 390 households reported this reason). The releases of new hybrid varieties with suitable maturation period, however, should have also contributed to the increase. Our survey shows that the actual growing time in 1988 for early, middle, and late hybrids averaged, respectively, 111 days, 152 days, and 122 days, compared to 109 days, 149 days, and 125 days for conventional early, middle, and late rice.

As for government support for the adoption of F_1 hybrids, 21.2% of households reported that chemical fertilizers were used as an award when the hybrids were first promoted, and another 2.6% reported that hybrid seed was subsidized. Currently, none of the surveyed households reported seed subsidies and only 4.4% of the households reported having fertilizer supports.¹¹

The survey also asked about cooking quality of hybrid rice. The responses of the sample households indicate that this is no longer an is-

sue. Among the 319 households growing both hybrids and conventional rice, 48% reported eating conventional rice for daily meals, while the other 52% reported eating hybrid rice. As for the rice used for special occasions like entertaining guests and celebrating festivals, 25.4% reported to prefer conventional rice, 39.8% to prefer hybrid rice, while the remaining 34.8% reported having no preference.

Of the 500 households surveyed, 495 devoted part of their land to rice. Detailed information on the number of households using hybrid and conventional seed in each of the five counties in 1988 is reported in table 4. While only a few (thirteen) among the 495 households planted hybrids in the early rice season, the majority of households adopted hybrid seed either in the middle season if only one crop of rice is grown or in the late season if two crops of rice are grown each year. A substantial portion of the households in each county planted both hybrid and conventional rice in a single crop season. This practice will enable us to investigate not only the problem of dichotomous choice but also the optimal allocation of land to the production of hybrids.

The reason for early hybrids not being adopted by most households may arise from the fact that the yield advantage of early hybrids with suitable maturation period was not significant compared to early conventional varieties. Table 5 reports the means and standard deviations of inputs and outputs for hybrid and conventional rice from the samples. While the mean yields of hybrids were significantly higher than those of conventional varieties for middle and late rice, the difference is not statistically significant for early rice. Hybrid rice's yield advantages are partly offset by added requirements for chemical inputs and more expenditures for seed. As indicated in table 5, significantly more chemical inputs (chemical fertilizers and pesticides) were devoted to hybrids as compared to conventional varieties. Also, although cultivation of hybrids

¹¹ This response may underreport the government's function in the promotion of F_1 hybrid rice, as government supports are mainly given to the seed research and distribution system and not directly to individual households.

Table 4. The Adoption of Hybrid and Conventional Rice in 1988

	Tiaojian (N = 100)	Xiangxiang (N = 100)	Nanxian (N = 97)	Anxiang (N = 99)	Zhijiang (N = 99)
Early Rice					
Hybrid	4	7	0	0	2
Conventional	98	98	92	98	6
Both	2	5	0	0	0
Middle Rice					
Hybrid	0	1	8	8	99
Conventional	0	0	11	9	14
Both	0	0	0	2	14
Late Rice					
Hybrid	79	67	63	90	9
Conventional	35	49	78	51	0
Both	14	18	46	43	0

required only about one-third to one-fourth as much seed as conventional varieties, the price of hybrid seed is fixed as ten times that of conventional seed. Therefore, the advantage of growing hybrids compared to conventional varieties depends largely on the prices of chemical inputs and seed, even for middle and late rice. However, hybrid rice does not require more labor input than conventional rice.

Empirical Analysis

Functional Form Specification

In the theoretical section, it was argued that the probability as well as the intensity of adopting F_1 hybrid seeds in rice production for a randomly selected household are functions of the

household head's education, and other variables representing the household's characteristics and the economic environment. It is clear from the theoretical model that the functions for the dichotomous choice and for the optimal adoption decision should have the same set of explanatory variables. Table 6 presents the definitions, means, and standard deviations of the variables which will be used as regressors in the empirical analysis. Variables 1 to 4 are county dummies that represent some county-specific characteristics, which affect the adoption decision but are not observable to researchers.¹²

¹² The county dummies capture the effect of differences in a county's topography, rainfall, temperature, and other physical variables on a household's adoption decision. Because hybrid seed production, distribution, and extension services are organized by each county government, the county dummies will also capture the effect of differences in these government-provided services.

Table 5. Means and Standard Deviation—Hybrid and Conventional Rice

	Early Rice		Middle Rice		Late Rice	
	Conv. (N = 392)	Hybrid (N = 13)	Conv. (N = 34)	Hybrid (N = 116)	Conv. (N = 213)	Hybrid (N = 308)
Seed (kg/mu)	11.6 (3.7)	3.1 (2.2)***	6.9 (3.1)	2.0 (0.9)***	6.2 (3.3)	2.0 (1.1)***
Fertilizer (Yuan/mu)	21.0 (9.4)	27.5 (7.9)*	14.9 (8.4)	18.9 (10.4)*	22.7 (13.2)	24.7 (10.7)*
Pesticide (Yuan/mu)	5.1 (4.7)	7.3 (2.7)	4.1 (3.5)	6.2 (4.1)*	5.7 (3.9)	7.3 (5.4)**
Labor (day/mu)	15.3 (6.4)	17.0 (3.8)	22.3 (13.6)	20.3 (7.4)	13.9 (5.8)	14.4 (5.4)
Draft animals (day/mu)	1.5 (1.1)	1.9 (1.0)	3.0 (3.0)	3.7 (2.1)	1.1 (1.1)	1.1 (.7)
Machine (day/mu)	.5 (.8)	1.3 (1.4)**	1.4 (6.9)	.06 (.31)*	.6 (.8)	.6 (.9)
Rice output (kg/mu)	352.7 (97.3)	385.4 (172.1)	270.0 (117.8)	432.5 (124.2)***	323.6 (90.2)	386.6 (85.3)***
Straw (kg/mu)	206.9 (109.9)	293.1 (57.4)**	211.9 (92.4)	295.9 (89.6)***	240.8 (95.7)	296.1 (98.1)***

Note: *, **, and ***, indicate, respectively, that the means are significantly different at the 5%, 1%, and 0.1% level of confidence.

Table 6. Variable Definitions, Means, and Standard Deviations

Dependent variables:		
(1) <i>AHR</i>	Adoption dichotomous = 1 if middle or late hybrid rice is adopted, 0 otherwise	.82 (.39)
(2) <i>PHR</i>	Percentage of middle and late rice area grown with hybrid seeds	.69 (.40)
Independent variables:		
(1) <i>C2</i>	County dummy variable = 1 if Nanxian County, 0 otherwise	.20 (.40)
(2) <i>C3</i>	County dummy variable = 1 if Anxiang County, 0 otherwise	.20 (.40)
(3) <i>C4</i>	County dummy variable = 1 if Zhijiang County, 0 otherwise	.20 (.40)
(4) <i>C5</i>	County dummy variable = 1 if Xiangxiang County, 0 otherwise	.20 (.40)
(5) <i>Ps</i>	Price of hybrid rice seed (Yuan/kg)	4.07 (.63)
(6) <i>Pf</i>	Price of chemical fertilizer (Yuan/kg)	.41 (.49)
(7) <i>Pp</i>	Price of pesticide (Yuan/kg)	10.06 (2.49)
(8) <i>LAND</i>	Land area cultivated (in mu, 1 ha = 15 mu)	6.46 (2.97)
(9) <i>LABOR</i>	Number of adults	4.39 (1.22)
(10) <i>CAPITAL</i>	Value of capital equipment (in Yuan)	459.22 (653.24)
(11) <i>CR</i>	Credit dummy = 1, if formal credit used in the previous two years, 0 otherwise	.41 (.49)
(12) <i>Quota</i>	Rice procurement quota dummy variable = 1 if quota exists, 0 otherwise	.97 (.16)
(13) <i>Avedu</i>	Average education level of other adult household members (in years)	5.18 (2.27)
(14) <i>Job</i>	Job dummy of household head, = 1 if nonfarm, 0 if farm	.10 (.30)
(15) <i>Sex</i>	Sex of household head, = 1 if female, 0 otherwise	.04 (.19)
(16) <i>Agryrs</i>	Household head's experience in agriculture (in years)	23.66 (11.61)
(17) <i>Eduhead</i>	Education level of household head (in years)	5.49 (2.58)

Variables 5 to 7 are price variables, representing the economic environment. Theoretically, the prices of seed, fertilizer, and pesticide relevant to the decision to adopt hybrid seeds should be expected prices. Because data on expected prices are not available, the actual prices a household paid were used as proxies for expected prices.¹³ If significant differences exist between expected prices and actual prices, the estimates of parameters may be biased. The price of rice is not included in the list because no cross-sectional variation in rice price existed. Because the information on wage, rent, and interest rate are not available, wage, rent, and interest rate are not included either.

¹³ The prices are derived from each household's actual expenditures on seeds, fertilizers, and pesticides, dividing by the quantities of seeds, fertilizers, and pesticides. Because some portion of each of the inputs was rationed and some was purchased from markets, the derived prices are the average prices paid by a household.

Variables 8 to 13 represent the household-specific characteristics, including endowments in landholding, labor, and capital, dummies for credit availability and state rice procurement quota, and the average education level of adult household members (excluding the household head).¹⁴ The last group of variables was the household head's personal characteristics, including dummies for job type and sex, years of experience in agriculture, and years of education.

As for the dependent variables in the analysis, only the adoption decision with respect to middle rice and late rice is considered. Early rice is excluded because the data indicate that adapted

¹⁴ Because information on the availability of credit does not exist, the incidence of taking formal credit in the previous two years was used as a proxy for credit availability. However, this proxy may not reflect actual credit availability because the absence of borrowing may indicate that a household has enough funds of its own.

early hybrids are not yet available. In the dichotomous choice model, a household is considered as an adopter if it grew either middle hybrids or late hybrids. In the model of optimal adoption decision, the dependent variable is the percentage of total rice acreage planted with hybrids in both the middle and late rice seasons.¹⁵

The theoretical model suggests that if the functional forms of $D(E)$ and $C_1(0|e_i, Z_i, E)$ are linear, and the unobservable variable ϵ_i is identically and independently distributed as the normal distribution over the population, then probit is the appropriate method for estimating the unknown parameters in the dichotomous choice model. Similarly, if the functional form of $r(e_i|e_i, Z_i, E)$ is linear in the explanatory variables and ϵ_i has a normal distribution, then the two-limit tobit is the appropriate method for estimating the unknown parameters in the optimal adoption rate equation. Therefore, the probit and tobit model will be applied to estimate the following function:

$$(14) \quad X = \alpha_1 + \alpha_2 C2 + \alpha_3 C3 + \alpha_4 C4 + \alpha_5 C5 + \beta_1 LnFs + \beta_2 LnPf + \beta_3 LnPp \\ + \gamma_1 LnLand + \gamma_2 LnLabor + \gamma_3 LnCapital + \gamma_4 CR + \gamma_5 Quota + \gamma_6 Avedu \\ + \delta_1 Job + \delta_2 Sex + \delta_3 Agryrs + \delta_4 Eduhead + \epsilon_x; \\ X = AHR \text{ (dummy variable for the dichotomous choice) or PHR} \\ \text{(percentage of acreage planted with hybrids).}$$

The price variables and input endowments are the logarithms of their respective quantities. This is because under the assumption of Cobb-Douglas production functions, the logarithmic normalized production function is a linear function in the natural logarithms of the normalized prices and the fixed inputs, and is also linear in the characteristic variables (Jamison and Lau, p. 203).

Empirical Results

The empirical results are presented in table 7. The estimates of coefficients indicate that a household head's education level has a positive effect both on the probability of adoption and the intensity of adoption of F_1 hybrids, as predicted by the theoretical model. The estimates are different from zero at a .1% level of significance in the dichotomous choice model and at a 5% level of significance in the optimal adoption decision model. This evidence gives support in a Chinese context to Schultz's thesis about

the role of education in decisions about the adoption of new technology.

As for the effects of other independent variables on the probability of adopting hybrid seed, the estimates in column (1) indicate that a household head's number of years of experiences in agriculture, the existence of government procurement quota, and the size of a farm's cultivated land also have significantly positive impacts, and the price of hybrid seed has a significant negative impact. Other economic and household characteristic variables do not have a significant effect. The positive impact of farm size on probability of adoption may arise from the economies of scale in acquiring information, credit, and/or hybrid seeds. As for the optimal adoption decision, column (2) of table 7 indicates that, besides a household head's education level and county dummies, a household's capital endowment is the only variable to have significant effect (at 10% level). This result im-

plies that higher investment in fixed inputs increases the proportion of land used in hybrids.

Concluding Comments

This paper has focused on the role of education in a farm household's decision about whether to adopt and the optimal intensity of adoption of F_1 hybrid rice in China. A simple behavioral model was developed in which adoption of new technology was treated as a portfolio selection problem, and the implications of the model were tested with data collected from a sample of 500 households from five counties in Hunan Province. The empirical results are consistent with the implications of the role of education in the theoretical model: A household head's level of education has positive and statistically significant effects on the household's probability and intensity of adopting F_1 hybrid seed.

Because technological change is the main force in agricultural development, the evidence in the paper supports arguments for increasing state investments in rural education in order to facilitate technological change in agriculture. In addition, a farm's size has a positive effect on the deci-

¹⁵ Most households grow either middle or late rice, but not both.

Table 7. Probit Estimates for Dichotomous Adoption and Two-Limit Tobit Estimates for Adoption Rate of Hybrid Seed

Dependent Variable:	Probit AHR (1)	Two-limit Tobit PHR (2)
Independent Variable:		
Constant	-3.40 (2.22)	-.81 (.73)
C2	-.43 (1.68)	-.38 (2.26)
C3	-.56 (1.37)	-.85 (3.15)
C4	.57 (1.56)	-.15 (.60)
C5	5.69 (.02)	.85 (2.82)
Ps	-.95 (1.71)*	.14 (.34)
Pf	.21 (.22)	-.20 (.35)
Pp	.47 (1.45)	.21 (.86)
LAND	.74 (2.91)***	.040 (.24)
LABOR	-.24 (.76)	-.173 (.80)
CAPITAL	.05 (.86)	.076 (1.88)*
CR	.17 (1.04)	-.08 (.70)
Quota	.69 (1.79)*	.44 (1.33)
Avedu	-.01 (.31)	.015 (.59)
Job	.14 (.52)	-.048 (.28)
Sex	.50 (1.21)	.43 (1.53)
Agryrs	.02 (1.78)*	.0025 (.44)
Eduhead	.16 (3.61)***	.052 (1.91)**
Log likelihood	-177.86	-426.73

Note: Figures in parentheses are absolute values of asymptotic t -statistics; *, **, and *** indicate that the estimates are significantly different from zero at .1, .05, and .01 level of confidence, respectively; definitions of variables are provided in table 6.

sion to adopt F_1 hybrids. This evidence suggests that the small farm size predominant following the household-based farming institutional reform may pose a restraint to technological change and thus supports the argument for further liberalization in land markets, in order to facilitate consolidation through land market transactions.¹⁶

¹⁶ Although collectivization can solve the problem of farm size, it will reduce a farmer's incentives for farming as well as adopting new technology (Lin 1988, 1991), therefore, collectivization should be excluded as an alternative for solving the problem of farm size.

[Received April 1990; final revision received October 1990.]

References

- Barker, Randolph, and Robert W. Herdt. *The Rice Economy of Asia*. Washington DC: Resources for the Future, 1985.
- Duraisamy, Palanigounder. "Human Capital and Adoption of Innovations in Agricultural Production: Indian Evidence." Economic Growth Center Disc. Pap. No. 577, Yale University, 1989.
- Feder, Gershon. "Farm Size, Risk Aversion and the Adoption of New Technology under Uncertainty." *Oxford Econ. Pap.* 32(1980):263-83.
- Feder, Gershon, Richard E. Just, and David Zilberman. "Adoption of Agricultural Innovations in Developing Countries: A Survey." *Econ. Develop. and Cultur. Change* 31(1985):255-98.
- He Gui-Ting, Zhu Xigang, and J. C. Flinn. "A Comparative Study of Economic Efficiency of Hybrid and Conventional Rice Production in Jiangsu Province, China." *Oryza* 24(1987a):285-96.
- . "Hybrid Seed Production in Jiangsu Province, China." *Oryza* 24(1987b):297-312.
- He Gui-Ting, Amanda Te, Zhu Xigang, S. L. Travers, Lai Xiufang, and R. W. Herdt. "The Economics of Hybrid Rice Production in China." *IRRI Res. Pap. Ser.*, No. 101, Dec. 1984.
- Jamison, Dean T., and Lawrence, J. Lau. *Farmer Education and Farm Efficiency*. Baltimore MD: Johns Hopkins University Press, 1982.
- Just, Richard E., and David Zilberman. "Stochastic Structure, Farm Size and Technology Adoption in Developing Agriculture." *Oxford Econ. Pap.* 35(1983):307-28.
- Lin, Justin Yifu. "The Household Responsibility System in China's agricultural Reform: A Theoretical and Empirical Study." *Econ. Develop. and Cultur. Change* 37(1988):199-224.
- . "The Household Responsibility System Reform and the Adoption of Hybrid Rice in China." *J. Develop. Econ.* (1991).
- Nelson, Richard R., and Edmund S. Phelps. "Investment in Humans, Technological Diffusion, and Economic Growth." *Amer. Econ. Rev.* 56(May 1966):69-75.
- Perkins, Dwight, and Shahid Yusuf. *Rural Development in China* (A World Bank Publication). Baltimore MD: Johns Hopkins University Press, 1984.
- Schultz, Theodore W. "The Value of the Ability to Deal with Disequilibria." *J. Econ. Lit.* 13(1975):827-46.
- . *Transforming Traditional Agriculture*. New Haven CT: Yale University, 1964.
- Statistical Bureau, Hunan Province. *Hunan Statistical Yearbook, 1989*. Beijing: China Statistics Press, 1989.
- State Statistical Bureau. *China Statistical Yearbook, 1989*. Beijing: China Statistics Press, 1989.
- Wiens, Thomas B. "Technological Change." *The Chinese Agricultural Economy*, ed R. Barker and R. Sinha. Boulder CO: Westview Press, 1982.

Incorporating Technological Change in Diffusion Models

Mary K. Knudson

The theoretical and empirical implications of a static and dynamic logistic diffusion model are compared. The dynamic model relaxes some assumptions of the static model by allowing for a flexible adoption ceiling, for changes in the technology, and for disadoption. Both models were used to estimate the diffusion of semi-dwarf wheat varieties in the United States. The dynamic model provides a better fit to the data and provides additional insights into the economic determinants of adoption. In particular, the importance of technological change in other areas (here, in fertilizer) on the diffusion path of an innovation was shown.

Key words: diffusion, semi-dwarf wheat varieties, technological change.

In constructing diffusion models, economists have faced a difficult trade-off. A host of probabilistic frameworks (such as the logistic, probit, and Gompertz; see Griliches 1957, Dixon) permit straightforward estimation of the diffusion trend of an innovation. Ease of estimation, however, comes at the cost of exogenously determined (and often theoretically unjustified) constraints on the rate of adoption and on the point at which the greatest rate of adoption occurs. For example, the logistic model requires the greatest rate of adoption to occur when 50% of the population has adopted the innovation. (For further criticisms of the static diffusion framework, see Mahajan and Peterson 1985.)

A number of the theoretically troublesome assumptions of the static diffusion model have been relaxed by Metcalfe and Gibbons. They present a dynamic diffusion framework which allows the inflection point of the diffusion curve to arise endogenously from the specifications of the model. Their model is of further interest because it depicts explicitly the interactive relationship between research and development and diffusion/adoption. Unfortunately, the data requirements for estimation of the Metcalfe and

Gibbons model are onerous. For example, input costs, including research costs, are required. However, firms are typically unwilling to release such input data.

This paper presents a method for overcoming data limitations in estimating dynamic diffusion models while incorporating additional variables to explain technological change. This methodology has potential application to a variety of diffusion cases, including the adoption of herbicide-resistant crop varieties and low-input agricultural practices. The technique is applied to estimating the diffusion of semi-dwarf wheat varieties (SDWVs). SDWVs are an important yield-increasing biological innovation which has substantially altered the agricultural practices of wheat growers in the United States. To emphasize the value of the dynamic approach, results from the dynamic model will be compared to estimates from a more traditional static diffusion model. As shall be shown, the dynamic model not only provides better estimates of the diffusion of semi-dwarf wheat, but also allows the researcher to determine the impact on the diffusion process of a variety of factors not incorporated within the traditional static model.

This paper is organized as follows. The first section presents the traditional static diffusion model, followed by Metcalfe and Gibbon's generalization to the dynamic case. The third section indicates how market data can be used to yield estimates for dynamic diffusion trends. The fourth section presents the results of a generalized least squares estimation of the traditional static diffusion model and of a dynamic diffu-

Mary K. Knudson is a visiting assistant professor, The Institute of Public Policy Studies, University of Michigan.

Work on this paper was supported by the Resources and Technology Division of the U.S. Department of Agriculture, Economic Research Service. The views expressed in this paper are the author's and do not necessarily reflect views or policies in this paper of the Economic Research Service.

The author would like to acknowledge Wiktor Adamowicz, Willis Peterson, Vern Ruttan, three anonymous referees, and, especially, Douglas Dion for their helpful suggestions. Remaining errors are the author's.

sion model for semi-dwarf wheat. Finally, some concluding remarks are made.

A Basic Diffusion Model

The basic diffusion process contains four components: the innovation, the social system in which the innovation is being adopted, channels of communication, and time (Metcalf and Gibbons, Mahajan and Peterson 1985, Bayer and Melone). The innovation component represents the new product being diffused. Inherent in the new product are characteristics that affect its diffusion.¹ A social system is comprised of individuals, organizations, or agencies and their adopting strategies. How a total group of adopters transmits information about an innovation depends on how heterogeneous and cohesive a group of adopters is. The ways that adopters pass on information about an innovation are the channels of communication. Finally, time is the period over which a social system adopts an innovation. Because these four components are present in every diffusion process, any realistic model of diffusion must incorporate them.

In addition, six basic assumptions underlie the diffusion model (See, for example, Mahajan and Peterson 1985.): (A1) The adoption decision is binary. An individual adopts or does not adopt. (A2) A fixed, finite ceiling N^M exists. (A3) The coefficient of diffusion (defined below) is fixed over time. (A4) The innovation is not modified once introduced, and its diffusion is independent from the diffusion of other innovations. (A5) One adoption is permitted per adopting unit and this decision cannot be rescinded. (A6) A social system's geographical boundaries stay constant over a diffusion process.

Incorporating the four components, and assuming (A1)–(A6), the basic diffusion model takes the following functional form (This discussion follows the notation found in Mahajan and Peterson 1985.):

$$(1) \quad dN(t)/dt = g(t) [N^M - N(t)],$$

where $dN(t)/dt$ is the rate of diffusion at time

t , (or, more simply, changes in adoption over time) where t begins at t_0 ; $g(t)$ is the coefficient of diffusion. This coefficient indicates how fast and to what degree adoption occurs. The higher the coefficient, the earlier and greater degree of adoption of an innovation.

The cumulative number of adopters at time t is $N(t)$, or

$$N(t) = \int_{t_0}^t n(t) dt,$$

and $n(t)$ is the number of adopters per time period t ; N^M is the maximum number of adopters in a social system over time t . The number of potential adopters left at time t is $N^M - N(t)$.

While it is clear how innovation and time fit into model (1), the entry of social system and channels of communication are not directly apparent. Their effect is found in the diffusion coefficient, $g(t)$. The specific form of $g(t)$ reflects the influence of the social system and the channels of communication on the diffusion process. For example, when a social system is heterogeneous, fragmented, and lacks communication, $g(t) = b$, where b represents the natural rate of diffusion, i.e., that rate which occurs without any other influence. The resulting diffusion model, $dN(t)/dt = b[N^M - N(t)]$, is called the external influence model. However, when the social system is homogeneous and communicative, $g(t) = bN(t)$, where b still is the natural rate of adoption and $N(t)$ is the impact adopters have on the remaining number of non-adopters, $N^M - N(t)$. The resulting diffusion model,

$$(2) \quad dN(t)/dt = bN(t) (N^M - N(t)),$$

is known both as an internal influence model and as the logistic model; it was used to estimate the diffusion of hybrid corn in Griliches' (1957) seminal piece.

The logistic model imposes an S-shaped, symmetric diffusion trend with a maximum diffusion rate occurring when 50% of the potential cumulative adopters have adopted.² It is based on the premise that diffusion occurs through interpersonal contacts among a group of homogeneous adopters (Mansfield). But not all diffusion models require symmetry around a 50% inflection point. For example, the Gompertz model, $dN(t)/dt = bN(t) (\log N^M - N(t))$, im-

¹ Economists sometimes classify innovations into process and product. A process innovation is an input to a production process. A product innovation is an end product. Generally, product innovations are easier to assess than process innovations because their performance can be measured directly. However, many innovations fall into both categories. That is, one firm's product innovation, say a seed variety, may become a process innovation for another firm. For the purposes of this paper, what category the innovation falls into does not matter as long as it is measurable. For simplicity, product is used.

² For proofs of the logistic model's inflection point and its symmetric diffusion trend, see Mahajan and Peterson 1985, pp. 27–28.

poses an asymmetric trend with the maximum diffusion rate occurring when 37% of the potential cumulative adopters have adopted.³ The assumption here is that although adopters are homogenous, early adopters are relatively more cohesive than middle and late adopters; therefore, they adopt at a faster pace. Dixon used the Gompertz and traditional logistic models to reestimate the diffusion of hybrid corn. He found the Gompertz model was a more accurate measure of hybrid corn diffusion.

Because farmers within a region or country often have similar production practices (homogenous) and communicate well with each other (cohesive), it makes sense to use a traditional logistic model to estimate the diffusion of an agricultural innovation (Mahajan and Peterson 1985). Hence, in this paper, the traditional logistic model is used to estimate the diffusion of semidwarf wheat. Because this model will be compared to a dynamic model, the traditional logistic model will hereafter be called the static logistic model.

Dynamic Diffusion Models

Static diffusion models work best when the adoption process being modeled satisfies the assumptions (A1)–(A6) listed above, and when the inflection point for the adoption trend occurs at the same point as the inflection point for the functional forms listed above. For many applications, however, static diffusion models are open to two objections. First, in many cases there will be no rationale *ex ante* for assuming that diffusion follows a particular trend. Second, the assumption of a fixed ceiling on the adopting population is unrealistic in most economic contexts. For example, the potential number of adopters of a biological innovation will vary depending upon the availability of the innovation,

which itself is a result of the profit-maximizing efforts of firms.

Models are needed that allow more flexibility with regard to the inflection and symmetry points (Mahajan and Peterson 1985, Knudson). Dynamic models allow the determinants of diffusion to change every time period, and, hence, may more accurately measure the rate of adoption than a static model. For example, as the real price of an innovation decreases and stabilizes, an innovation becomes more attractive and is adopted more rapidly. A dynamic model could capture this change; a static model could not.

In addition, as a result of its flexible form, dynamic models can include more variables that affect diffusion and therefore measure more directly these factors' impact. For instance, the impact of such factors as prices or profit can be included in the diffusion model. Conceptually, then, a diffusion trend becomes an envelope of individual, distinct diffusion trends that a set of distinct variables define in each time period (Chow; Mahajan and Peterson 1978, 1985; Metcalfe and Gibbons; Stoneman).

Studies by Mahajan and Peterson (1978) and by Metcalfe and Gibbons were used to develop such a dynamic model. In particular, the Mahajan and Peterson study established an approach for modifying static models. They began by relaxing assumption (A2), which states that N^M is fixed and finite, claiming that (A2) is inconsistent with theory and practice. To make N^M dynamic, Mahajan and Peterson defined N^M as a function $f(s(t))$, where $s(t)$ is a vector of "(potentially) all relevant exogenous and endogenous factors, both controllable and uncontrollable, affecting $N^M(t)$ " (p. 1590). How well $f(s(t))$ represents $N^M(t)$ depends on the accuracy in estimating $s(t)$ and the closeness of $s(t)$ or $f(s(t))$ in measuring $N^M(t)$. Hence, by redefining $s(t)$ or $f(s(t))$, this approach can be tailored to measure the diffusion of distinctly different innovations.

Metcalfe and Gibbons relaxed assumptions (A2)–(A4) to build a dynamic model that is "based upon the simultaneous development of market demand and the accumulation of productive capacity" (p. 5). They take a static logistic model and make $N^M(t)$ a function of a simple market demand equation, $c - ap(t)$, where $p(t)$ is prices paid by the consumer at time t . Their $p(t)$ is determined endogenously within an industry; hence, it represents the equilibrium price (reflecting the growth rate capacity and profit rate of a firm). Their model takes the final form;

³ A Gompertz model takes the form, $dN(t)/dt = bN(t) [\log N^M - \log N(t)]$, where b , and $N(t)$ and N^M are as defined in the text. The following is a proof that the Gompertz model has a maximum diffusion rate when 37% of the potential adopters adopt and is asymmetric.

Let N be expressed as a fraction of N^M , so $dN(t)/dt = bN(t) (\log N^M - \log N(t))$ becomes $(dN(t)/dt) (1/N^M) = b(N(t)/N^M) (\log(N^M/N^M) - \log(N(t)/N^M))$. Let $F(t) = N(t)/N^M$, and ignore the subscript t for convenience; $(dN(t)/dt) (1/N^M)$ becomes $dF/dt = bF(\log 1 - \log F) = bF(\log(1/F))$.

To find the maximum growth rate, F^M , differentiate with respect to F ; $(dF/dt)/dF = b \log(1/F) - bF^M(1/F^M) = 0$; rearranging, $b \log(1/F^M) = b \log(1/F^M) = 1$ or $F^M = 0.37$. The maximum diffusion rate occurs when 37% of the potential adopters have adopted and is skewed to the left, making the diffusion trend asymmetric.

$$(3) \quad dN(t)/dt = g(t) [N^M(t) - N(t)],$$

where $g(t)$ is a function of technology, the natural rate of diffusion, and the capacity for a firm to invest in research and development.

The approach utilized by Metcalfe and Gibbons relaxes many of the basic assumptions of the diffusion model listed in the preceding section. By allowing the adoption ceiling to be a function of time (as well as technology and the parameters of equilibrium demand and supply), the assumption of a fixed adoption ceiling (A2) is relaxed. All three components $g(t)$ listed above may change every time period; hence (A3) (the assumption that the diffusion coefficient is fixed) is relaxed. (A4) is relaxed as the model allows directly for improvements in the innovation and for improvements in the means of producing the innovation. (For further details of model (3); see Metcalfe and Gibbons and Metcalfe.)

Modeling the Diffusion of Semi-Dwarf Wheat

Metcalfe and Gibbons present their model as a means of incorporating the full effects of the invention side in the diffusion process. Unfortunately, the cost data required for the profit rate equation are virtually impossible to obtain. Similarly, revenue for a research project is difficult to obtain, and as a result one cannot estimate marginal revenue to approximate marginal cost. Furthermore, not many proxies for research costs are available. The Metcalfe and Gibbons model, while theoretically appealing, is difficult to estimate empirically.

A model overcoming these data limitations was developed to estimate the diffusion of semi-dwarf wheat varieties (SDWV) across the U.S. Let the maximum number of adopters be a function of a wheat supply function, $N^M(t) = f(s(t)) = y(t)$, where y is the supply of wheat from semi-dwarf wheat seeds that farmers produce in time t . More specifically,

$$(4) \quad y(t) = f(pr(t), pp(t), pf(t)) \\ = c + a^1 pr(t-1) + a^2 pp(t-1) \\ + a^3 pf(t-1),$$

where $pr(t-1)$ is the price farmers receive for a bushel of grain in year $t-1$, $pp(t-1)$ is the price paid for SDWV seed in time $t-1$, and $pf(t-1)$ is the price paid for fertilizer in time $t-1$.

Two factors stand out in this functional form (3). First, the price variables are lagged one year.

Because a substantial lag exists between the time when production decisions are made to when production ends, price variables are often lagged. In this case, the hypothesis that a producer's expectation of prices is based on relatively recent experience is used (Tomek and Robinson). Second, the price variables are not deflated by a common deflator. Instead, the price variable that would have been the deflector is represented as a separate variable. In this case, the effect of fertilizer prices is measured.

Measuring the impact of fertilizer prices as an independent variable makes sense for two reasons. First, for the innovation studied here (SDWV), high yields were obtained through the use of heavier fertilizer application. Therefore, the use of fertilizer as a deflator may bias the regression results (Tomek and Robinson). Second, because fertilizer prices fluctuated greatly in the 1970s, separating prices received and other input prices from fertilizer prices allows for more accurate measurements of the impact of these variables on supply (Meilke).

The main difference between this model of diffusion and that of Metcalfe and Gibbons is that prices are now exogenous. However, because the supply function depends on factors that affect a farmer's decision to adopt, a supply function still connects the stages of research and development and diffusion/adoption. For example, suppose high fertilizer prices deter the adoption of SDWV. A firm may use this information to develop SDWVs that are less dependent on fertilizer. Furthermore, as prices favor a particular production process, such as SDWV and increased fertilizer use, farm producers may use these inputs to utilize a profitable opportunity (Schumpeter). Work by Peterson, and Timmer and Falcon show for various commodities that as the real price of the agricultural product increases, its supply also increases. Hence, a supply function similar to that found in equation (4) also captures the ceiling of SDWV adoption.⁴

Substituting (4) for N^M in equation (1) yields

$$(5) \quad dN(t)/dt = bN(t) (c + a^1 pr(t-1) \\ + a^2 pp(t-1) + a^3 pf(t-1) - N(t)),$$

where $dN(t)/dt$ is the rate of diffusion in time t , $N(t)$ is the number of adopters in time t , b is the natural rate of adoption, and $c, a^1, a^2, a^3, pr(t$

⁴ The demand elasticity is assumed not equal to zero since in that case firms would not respond to the lowered price of the innovation. Therefore, technological change would have no effect on the firm.

-1), $pp(t-1)$, and $pf(t-1)$ are defined above. In equation (5), as in the Mahajan and Peterson model, the maximum number of adopters does not remain fixed over time; therefore, (A2) is relaxed. Similar to the Metcalfe and Gibbons model, the innovation is allowed to change over the diffusion process through the supply function found in equation (4); hence, (A4) is relaxed.

Estimating the Diffusion of SDWV

The parameters for the static logistic model (2) and the dynamic logistic model (5) are estimated by rewriting each equation in its discrete analogue form and then applying generalized least squares (GLS) (Mahajan and Peterson 1985, Bass, Chow, Bayer and Melone, White).⁵ In terms of its discrete analogue form, $dN(t)/dt = N(t) - N(t-1)$. Substituting the discrete analogue form into the static logistic model (2), and multiplying the right-hand side of (2) out, (2) becomes

$$(6) \quad N(t) - N(t-1) = (bN^M)N(t) - bN(t)^2,$$

where $bN^M = A$, $-b = E$, and $N(t)$ and $N(t)^2$ are the same as they were [in (2)]. Equation (6) becomes by substitution,

$$(7) \quad N(t) - N(t-1) = AN(t) + EN(t)^2.$$

Doing the same for the dynamic logistic model, equation (5) becomes

$$(8) \quad N(t) - N(t-1) = (bc)N(t) + (ba^1)N(t)pr(t-1) + (ba^2)N(t)pp(t-1) + (ba^3)N(t)pf(t-1) - (b)N(t)^2.$$

Let $A = bc$, $B = ba^1$, $C = ba^2$, $D = ba^3$, and $E = -b$, and substitute, A , B , C , D , and E for these terms in (8). Now, (8) becomes

$$(9) \quad N(t) - N(t-1) = AN(t) + BN(t)pr(t-1) + CN(t)pp(t-1) + DN(t)pf(t-1) + EN(t)^2.$$

One can now apply GLS to both equations (7) and (9).

For both models, the A coefficient is expected to be positive and the E coefficient negative. The greater the increase in the number of adopters, $N(t)$, between two time periods, the greater the

diffusion rate, $N(t) - N(t-1)$. However, as the number of adopters increases, the number of potential adopters decreases. Adoption continues to increase but at a decreasing rate.

For the dynamic logistic model, a positive B coefficient is expected and negative C and D coefficients are expected. Producers supply more of a commodity as its real market price increases. Hence, as prices increase, so does adoption and the diffusion rate. However, as an input price decreases, assuming everything else remains constant, the profitability of using that input increases, causing its adoption to increase.

Variables Used

Both the static and dynamic logistic models use the variable $N(t)$. In addition, the dynamic model also uses the variables $pr(t-1)$, $pp(t-1)$, and $pf(t-1)$. $N(t)$ is measured by the percentage of total wheat land planted to SDWV's. This measure is a flexible variable, capable of capturing disadoption trends. This allows one to take advantage of relaxing assumption (A5).

An alternative measure of $N(t)$ is the area of land planted with SDWV's. One advantage of using the relative measure is to maintain consistency with other studies that measured the diffusion of a new crop variety. Griliches (1957) and Dixon used percentage of corn land planted to hybrid corn to measure hybrid corn adoption. A second justification for using the relative measure is that this study addresses the extent to which wheat farmers use SDWVs as opposed to traditional varieties. Because the only commercial alternative for wheat farmers is either the traditional or the SDW option, the relative proportion of wheat land going to SDW allows one to track the adoption of SDWV over traditional varieties within the adoption population of interest, i.e., wheat farmers. By contrast, using the absolute measure gives no insight into the adopting behavior of wheat farmers because increases in the total land planted with SDWVs need not reflect displacement of the more traditional varieties.⁶

Data on the percent of SDWV acreage relative to total wheat acreage are collected from a

⁵ The Goldfeld-Quandt test for heteroskedasticity was statistically significant and was corrected following the methodology of White.

⁶ As a check, the model was rerun using the absolute acreage measure. The results were confusing. For example, the coefficients for $N(t)$ imply that the adoption occurred first in regions where the SDWV technology appeared last and conflicts with evidence (such as the data from the Wheat Marketing Survey) that shows otherwise. This result may reflect the erratic wheat acreages during the 1960s and 1970s.

wheat marketing survey by U.S. Department of Agriculture (USDA) every five years. These data are aggregated so the unit of analysis is on a per state basis. The survey is run across forty-two states. Data on SDWV acreage allotment are available for forty-one of these states. Florida is excluded because of lack of data for many years on their wheat plantings.

The survey years for which SDWV data are available are (a) 1959, 1964, 1969, 1974, 1979, and 1984 for Idaho, Montana, Nevada, Oregon, and Washington; and (b) 1964, 1969, 1974, 1979, and 1984 for the remaining thirty-six states.

The three price variables in the model (Pr , Pp , and Pf) are measured as follows: Pr , expected wheat prices received, is set equal to the average annual price received per bushel of wheat. ($N(t - 1)$ refers to the preceding wheat survey year. However, the price variables are lagged only one year.) The price received "allows for unredeemed loans and government purchases values at the average loan and purchase rate" (USDA ERS, p. 24). Expected prices paid for seed, Pp , equals the average annual price paid for seed. The USDA publishes seasonal average wheat price data. Finally, Pf , expected prices paid for fertilizer, is measured by total plant nutrient expenditures (million dollars) ($Pf * Qf$, where Pf is price of plant nutrients and Qf is quantity of plant nutrients) divided by the quantity of plant nutrients consumed on the national level (1,000 tons), i.e., $Pf * Qf / Qf = Pf$. This price accounts for technological improvements in fertilizer. As improvements occur, a unit of fertilizer will produce more nutrients, hence, reducing the real price of fertilizer. Because prices using plant nutrients are adjusted for quality and USDA fertilizer prices are not, USDA fertilizer prices are biased upwards. Therefore, this study used this measure ($Pf * Qf / Pf$) of fertilizer prices.

Table 2. Characteristics for States within Each Group

Group	Wheat Market Class
G1	White
G2	Hard Red Spring
G3	Hard Red Winter
G4	Soft Red Winter
G5	Soft Red Winter

Cross-Sectional and Time-Series Data

Because of the lack of time-series data on a per state basis and the abundance of cross-sectional data, equations (7) and (9) were estimated using cross-sectional and time-series data. The level of analysis using cross-sectional and time-series data is the individual as opposed to a time point. For this study, the movement of SDWV adoption in one region will be described relative to its movement in another region. The advantage is that all adopters' behavior will be directly comparable.

Equations (7) and (9) must be reformulated for use in cross-sectional and time-series analysis using a slope dummy variable framework. The forty-one states are divided into five groups based on their location and similarity in wheat and growing conditions. See tables 1 and 2 for a listing of the groups and their characteristics.

A base group was selected to provide a means of comparing SDWV diffusion rates between regions. The other groups were formulated using dummy variables. The model measures the diffusion of SDWV for the base group and how the other groups differ directly from the base group. Therefore, the base group must be different enough from the rest of the groups so that statistical differences between groups can be picked up. In this study, a good base group would

Table 1. State Groups

Group One (G1) (West)	Group Two (G2) (Upper Midwest)	Group Three (G3) (Plains)	Group Four (G4) (South)	Group Five (G5) (East)
Arizona	Michigan	Colorado	Alabama	Delaware
California	Minnesota	Illinois	Arkansas	Indiana
Idaho	Montana	Iowa	Georgia	Kentucky
Nevada	New York	Kansas	Louisiana	Maryland
New Mexico	North Dakota	Missouri	Mississippi	New Jersey
Oregon	South Dakota	Nebraska	North Carolina	Ohio
Utah	Wisconsin	Oklahoma	South Carolina	Pennsylvania
Washington		Texas	Tennessee	Virginia
		Wyoming		West Virginia

be one where either high or low SDWV diffusion occurred. Since SDWV diffusion was most rapid in Group 1, the West, this region was chosen as the base group.⁷

Results and Discussion

The results of the static logistic model (7) and the dynamic logistic model (9) are presented in tables 3 and 4, respectively. In both tables, the reported coefficients equal the coefficients from the base group plus the coefficient from the dummy variable for that particular variable in that region. The standard errors and *T*-statistics, however, are for the test that the coefficient for that region is significantly different from the coefficient for the base region (the West). Only the *T*-statistic for the West region is appropriate for determining whether the coefficient is significantly different from zero.

Overall, the dynamic model provides a better fit to the data than the static model. Adjusting

for additional explanatory variables in the dynamic model increased the *R*-squared measure from 0.45 to 0.67. In addition, the pattern of diffusion in the two models is quite similar. Considering the coefficients on *N* and *N*², both models show similar patterns of signs for the coefficients. In the West and Upper Midwest, increases in the number of cumulative adopters leads to increases in the diffusion rate, although the increases occur at a diminishing rate for larger numbers of cumulative adopters. This deceleration is not observed in any of the other three regions, where both the static and the dynamic diffusion models show a positive coefficient for the *N* and *N*² coefficients. However, these positive coefficients for *N*² for the Plains, the South, and the East may indicate that SDWV have not yet fully diffused in these regions.

Analysis of the value of wheat relative to total farm production in each region substantiates the patterns revealed in the *N* and *N*² coefficients. As noted in Griliches (1958), the higher the total value of a crop, the faster the acceptance rate. Wheat has the highest value in the West, the upper Midwest, and the Plains. The South and the East lag behind. This relationship suggests that adoption of SDWVs has progressed most fully in the West and upper Midwest, and less so in the other regions. If true, this result suggests that future empirical analysis of the diffusion of SDWVs in the Plains, South, and East should yield the expected negative coefficient *N*², as the regions become saturated and increases in land devoted to SDWVs decelerates.

⁷ Each of the groups act as one unit. The states within each group have fairly homogenous SDWV diffusion rates. Some states within a group may lag behind. But, because of the spillover effect, close locality, and similar market classes of wheat grown, it is assumed SDWV technology was similarly transferred among these states.

In order to check the homogeneity among states within each group, each group was run by itself in a cross-section and time-series logistic framework. The base states were Oregon, New York, Oklahoma, South Carolina, and Virginia. With the exception of Kansas, Texas, Indiana, Pennsylvania, North Carolina, and Mississippi, the states, within their respective groups, were not significantly different from the base group. Other groupings were also tested, but only this division gave strong results within and between regions.

Table 3. Results from Estimating a Static Logistic Using Only the Number of Adopters

Variable	Coefficient	Standard Error	<i>T</i> -Statistic
<u>West (G1)</u>			
<i>N</i>	0.8204025	0.2500690	3.280704
<i>N</i> ²	-0.007712485	0.003044787	-2.533013
<u>Upper Midwest (G2)</u>			
<i>N</i>	0.4267331	0.3326181	1.183548
<i>N</i> ²	-0.001141102	-0.005183085	1.267852
<u>Plains (G3)</u>			
<i>N</i>	0.6785527	0.3200809	0.4431686
<i>N</i> ²	0.0007691705	0.00458963	1.848004
<u>South (G4)</u>			
<i>N</i>	0.2359555	0.3761646	1.5537
<i>N</i> ²	0.00471732	0.005530218	2.247616
<u>East (G5)</u>			
<i>N</i>	0.1846409	0.3420848	1.858491
<i>N</i> ²	0.0121731	0.00550287	3.71505
RBAR**2	.46		
D.F.	159		

Note: States and characteristics of each group are found in tables 1 and 2.

Table 4. Results from Estimating a Dynamic Logistic Using the Number of Adopters, Prices Received, Prices Paid, and Fertilizer Prices for Five Groups

Variable	Coefficient	Standard Error	T-Statistic
<u>West (G1)</u>			
<i>N</i>	1.179268	.2736722	4.309052
<i>Pr</i>	-.09329465	.05053714	-1.846061
<i>Pp</i>	.05736491	.06157601	.9316114
<i>Pf</i>	-5.787621	2.434321	-2.377509
<i>N</i> ²	-.002765783	.002829609	-.9774436
<u>Upper Midwest (G2)</u>			
<i>N</i>	1.076938	.3117690	-.3282208
<i>Pr</i>	.04508504	.06622319	2.089596
<i>Pp</i>	.4549232	.1012592	3.926146
<i>Pf</i>	-21.07764	3.468241	-4.408580
<i>N</i> ²	-.006138496	.003294418	-1.023766
<u>Plains (G3)</u>			
<i>N</i>	1.42598	1.301191	.1896052
<i>Pr</i>	-.3351066	.4511594	-.5359790
<i>Pp</i>	.0971373	.08452221	.4705556
<i>Pf</i>	-1.78292	4.063782	.9854614
<i>N</i> ²	.001267546	.004928605	.8183510
<u>South (G4)</u>			
<i>N</i>	1.251281	.3435780	.2095975
<i>Pr</i>	-.5080875	.1913784	-2.167397
<i>Pp</i>	-.1513808	.09975378	-2.092610
<i>Pf</i>	9.663808	4.728954	3.267410
<i>N</i> ²	.001882715	.004080817	1.139110
<u>East (G5)</u>			
<i>N</i>	.5686835	.3426880	1.781749
<i>Pr</i>	-.2944101	.2310009	-.8706264
<i>Pp</i>	-.01216012	.1352383	-.5140929
<i>Pf</i>	4.847369	7.445705	1.428339
<i>N</i> ²	.008456703	.004070585	2.756971
RBAR**2	.67		
D.F.	144		

¹ States and characteristics of each group are found in tables 1 and 2.

The price coefficients from the dynamic model are less enlightening. In particular, the signs on the coefficients for the prices received (*Pr*) and prices paid for seed (*Pp*) do not follow any clear pattern. There are four potential explanations for this failure.⁸ First, farmers may not be price responsive, at least to these two variables. Second, these price variables may have a significant effect but not in a way that the model estimated here can uncover. For example, prices paid for seed may interact with another variable

(such as farm income), and through this interactive effect may significantly affect adoption. Finally, coefficient estimates may be confused because of the lack of variation on the *Pr* and *Pp* variables. For example, this lack of variation may be attributed in the former case to government programs maintaining artificially high prices for wheat. Fourth, the estimated model suffers from collinearity, since the price variables are all multiplied by *N*, which is also an independent variable in the model.⁹

⁸ The dynamic model was run using long-term prices for all three variables (measured by a five-year lagged average). Long-term prices not only gave signs that followed less of a pattern than those obtained in the short-run prices case, but also had lower *t*-statistics. Thus, it is not the use of short-run output prices per se that is responsible for the pattern of coefficients on the price variables. Theoretical justification for the use of one-year lagged prices is presented in Tomek and Robinson.

⁹ The correlations between *N* and *NPr* and between *N* and *NPp* were roughly 0.9. The correlation between *N* and *NPf* was slightly less. Although there are techniques to correct for near multicollinearity, such as bayesian mixed estimation and adding data, neither were feasible here (because of the lack of a second data set). Results from models with multicollinearity are unbiased, although inefficient. This implies that multicollinearity cannot account for the switched signs on some of the price variables, although it may account for the small *t*-statistics.

On the other hand, fertilizer prices (P_f) have influenced the diffusion of SDWV, particularly in the West, the Upper Midwest, and the Plains (the three high value-of-wheat regions). For these three regions, the sign is correct, indicating that as the price of this input decreases, farmers purchased more SDWV seed. It is not surprising that fertilizer prices have a significant impact on SDWV diffusion. The improved yield capacity of SDWV over standard tall varieties was largely the result of its ability to withstand heavier heads brought about by increased fertilizer applications.

Fertilizer prices appear less important in explaining the behavior of the late-adopting regions of the South and the East. This result is consistent with Cochrane's treadmill theory (Cochrane 1958, 1986). According to Cochrane, as diffusion proceeds, the role of prices may diminish if the higher production dampens prices. While the initial increase in production because of the adoption of a new technology may not be enough to increase total supply, early adopters reap the benefits of new technologies through receiving old prices and increasing output per unit of input. However, middle and late adopters cause supply to increase appreciably. As a result, middle and late adopters do not benefit from an increase in the total value product over an increase in total costs.

Concluding Comments

This paper has presented a method by which diffusion models can incorporate dynamic factors such as changes in the maximum number of adopters and in the technology and the possibility of disadoption. The approach allows the researcher to avoid the extreme data burdens (in particular, the necessity for cost data) that arise in other dynamic diffusion models. This method was used to estimate the diffusion of semi-dwarf wheat varieties (SDWVs) in the United States from 1959–84. Comparisons with results from a static diffusion model show that the dynamic model provides a better fit to the data as well as offering insights into the economic determinants of adoption.

The results of this paper also indicate the importance of including other innovations that affect the diffusion or development of the innovation being studied. The pattern of adoption of SDWV's was affected to a considerable extent by changes in fertilizer prices, which themselves resulted from technological innovations

in the fertilizer industry. Incorporating these complex relationships may be necessary in order to explain satisfactorily the patterns of adoption of technological innovations. For example, a firm trying to market herbicide resistant varieties must consider the price of herbicide and other inputs as well as the management skills needed to assure the prompt adoption of these new varieties. Policy analysts may need to consider how price affects the adoption of new agricultural practices, such as low input agricultural practices.

This point is pertinent to many agricultural process innovations being developed today. Due to the poor prices of the 1980s, and the current concern over the environment, many agricultural companies are marketing production packages instead of just one input. Hence, they need models, such as the one developed in this paper, that allow for the impact of another innovation.

[Received April 1989; final revision received November 1990.]

References

- Bass, F. M. "A New Product Growth Model for Consumer Durables." *Manag. Sci.* 15(1969):215–27.
- Bayer, J., and N. Melone. "Predicting Acquisition and Adoption of Software Engineering Innovations." Work. Pap. No 42-86-87, Carnegie-Mellon University, 1987.
- Chow, G. C. "Technological Change and the Demand for Computers." *Amer. Econ. Rev.* 57(1967):1117–30.
- Cochrane, W. W. "A New Sheet of Music. How Kennedy's Farm Adviser Has Changed His Tune About Commodity Policy and Why." *CHOICES* 1(1986):11–15.
- . *Farm Prices. Myth and Reality*. Minneapolis: University of Minnesota Press, 1958.
- Dixon, R. "Hybrid Corn Revisited." *Econometrica* 48(1980):1451–61.
- Griliches, Z. "Hybrid Corn: An Exploration in the Economics of Technological Change." *Econometrica* 25(1957):501–22.
- . "Research Costs and Social Returns: Hybrid Corn and Related Innovations." *J. Polit. Econ.* 66(1958):419–31.
- Knudson, M. K. *The Invention and Diffusion of Two Competing Technologies: Semi-Dwarf and Hybrid Wheat*. Ph.D. thesis, University of Minnesota, 1988.
- Mahajan, V., and R. A. Peterson. "Innovation Diffusion in a Dynamic Potential Adopter Population." *Manage. Sci.* 24(1978):1589–97.
- . *Models for Innovation*. Beverly Hills CA: Sage Publications, 1985.
- Mansfield, E. "Technical Change and the Rate of Imitation." *Econometrica* 29(1961):741–66.
- Meilke, E. D. "Another Look at the Hog-Corn Ratio." *Amer. J. Agr. Econ.* 59(1977):216–19.

- Metcalf, J. S. "Impulse and Diffusion in the Study of Technical Change." *Futures* (1981):347-59.
- Metcalf, J. S., and M. Gibbons. "Industrial Policy and the Evolution of Technology." Paper presented at conference, *Technological Innovation and Production Structure: The Position of Italy*, Milan, Italy, 1983.
- Peterson, W. L. "International Farm Prices and the Social Cost of Cheap Food Policies." *Amer. J. Agr. Econ.* 61(1979):12-21.
- Schumpeter, J. *The Theory of Economic Development*. Cambridge MA: Harvard University Press, 1934.
- Stoneman, P. *The Economic Analysis of Technological Change*. New York: Oxford University Press, 1983.
- Timmer, C. P., and W. P. Falcon. "The Political Economy of Rice Production and Trade in Asia." *Agriculture in Development Theory*, ed. L. G. Reynold, pp. 373-410. New Haven CT: Yale University Press, 1975.
- Tomek, W. G., and K. L. Robinson. *Agricultural Product Prices. Second Edition*. Ithaca: Cornell University Press, 1981.
- U.S. Department of Agriculture, Economic Research Service. *State-Level Wheat Statistics, 1949-1988*. Statist. Bull. No. 779. Washington DC, 1989.
- White, H. "A Heteroskedasticity-Consistent Covariance Matrix Estimator and Direct Test of Heteroskedasticity." *Econometrica* 48(1980):817-38.

A Bayesian Approach to Explaining Sequential Adoption of Components of a Technological Package

Howard D. Leathers and Melinda Smale

Agricultural innovations are often promoted as a package—a new seed variety, a recommended fertilizer application, and other recommended cultivation practices. Nevertheless, many farmers adopt pieces of the package rather than the whole, in a sequential fashion. This paper presents a behavioral model which explains sequential adoption as a consequence of learning by adopting farmers. In order to learn more about the entire technological package, the farmer may adopt a part of the package. The model is shown to be consistent with observed patterns of sequential adoption.

Key words: Bayesian learning, diffusion of technology, technology adoption.

In U.S. agriculture before 1950, the “Green Revolution” of the 1960s, and in other regions since then, agricultural research institutions have developed and promoted innovations as packages of farming practices consisting of a seed variety, fertilizer application, and other cultivation techniques (Moseman, Ryan and Subrahmanyam, Lele and Goldsmith). Empirical evidence suggests, however, that farmers often adopt pieces of the package rather than the whole, in a sequence or “stepwise” fashion (Byerlee and Hesse de Polanco, Mann, Ryan and Subrahmanyam).

Economists working with agricultural research institutions have argued that sequential adoption may reflect rational choice by farmers, but the reasons for the decision are not clearly understood. Byerlee and Hesse de Polanco state that “because of capital scarcity and risk considerations, farmers are rarely in a position to adopt complete packages” (p. 519). Even when provided with advantageous credit and input delivery arrangements, “farmers . . . experiment with recommendations, often adopting them in stages rather than as a complete package” (Cummings, p. 24). In their review of literature

on agricultural innovations in developing countries, Feder, Just, and Zilberman conclude that “further work is needed to understand any order and regularity in such chain processes” (p. 288).

Explanations for sequential adoption have not yet been framed in a behavioral model, and most behavioral models have investigated the factors that affect the adoption of a single component or a package, rather than the pattern of component adoption (O’Mara; Hiebert; Lindner, Fischer, and Pardey; Lindner; Lindner and Fischer). Feder’s model is one of the first to explore how the interrelationship between recommended inputs influences a farmer’s adoption decision.

In this paper, we offer a Bayesian model that explains sequential adoption even when farmers are risk neutral and unconstrained in their expenditures. The model depicts a farmer who is uncertain whether the information she/he receives about a recommended innovation accurately describes her/his own prospects. Uncertainty is reduced but may not be resolved through experience. The model demonstrates that in order to learn more about the innovation, the farmer may choose to adopt a component rather than the package, even when the package is more profitable.

Empirical evidence of sequential adoption is reviewed in the next section, and several plausible explanations for sequential adoption are explored. Following the Cummings suggestion that farmers adopt components to “experiment”

Howard D. Leathers is an assistant professor, and Melinda Smale is a graduate student at the University of Maryland.

This research was funded in part by a grant from Centro Internacional de Mejoramiento de Maíz y Trigo. The views expressed are those of the authors.

Many helpful comments were received from Paul Heisey and Derek Byerlee. Remaining errors are those of the authors.

with recommendations, a Bayesian model of learning is presented which can explain sequential adoption. A simplified version of the model illustrates the main points and the implications of the model are discussed.

Evidence on and Explanations of Sequential Adoption

In a variety of settings, new technologies have been introduced as packages of innovations (Moseman, Lele and Goldsmith). Typically, the technological package is developed by interdisciplinary teams of plant breeders, agronomists, and soil scientists that seek to maximize yield by capturing the biological and chemical interactions among its components. In a rural environment where information dissemination costs are high, the package represents a "clear model to strive for in the longer run" and rewards early adopters with "dramatic yield increases to motivate others" (Mann, p. 75). Technological packages are designed to exploit interactions among the components. For example, yield increases from adoption of new seed may be further improved by applying appropriate levels of fertilizer and irrigation or modifying plant density.

The question to be addressed is: Since these interactions exist, why do farmers adopt pieces of the package at first (failing to take advantage of the interactions), and only later adopt the entire package? Empirical evidence of sequential adoption has been found in Turkey, Mexico, India, Pakistan, and Zimbabwe. (See Mann, Byerlee and Hesse de Polanco, Husain, Lowdermilk, Heisey et al., CIMMYT, Rohrbach, Ryan and Subrahmanyam). The empirical evidence shows the following: (a) For the most part, farmers choose to adopt inputs sequentially, adopting initially only a part of the package and subsequently adding components over time. (b) In some instances, farmers adopt a component and subsequently revert to the traditional agricultural practices. (c) Adoption patterns vary by agroecological zone (i.e., in one part of Pakistan, a new seed is typically adopted first; while in a different part of Pakistan, fertilizer is adopted first). (d) Adoption patterns vary within agroecological zones, among farmers facing similar markets and institutions.

The authors cited above have identified profitability, riskiness, uncertainty, lumpiness of investment, and institutional constraints as pos-

sible explanations for sequential adoption. A simple economic explanation of adoption behavior would suggest that a farmer will choose the most profitable technology and will continue to use that technology as long as relative prices remain unchanged. This simple model suggests: (i) Farmers operating under identical conditions will choose identical technological combinations, contradicting observation (d) above. (ii) As economic conditions change, farmers will switch from a component to the whole package in some cases, and from the whole package to a component in other cases, contradicting observation (a).

Ryan and Subrahmanyam postulated that "each part of the package might be looked upon by farmers as a less risky activity than the complete package in terms of what the farmer could lose if crop failure results" (p. A101). In other words, although expected profits may be maximized by adopting the package, expected utility may be maximized by adopting a component. This explains why a farmer chooses a part rather than the package, but not how the choice evolves over successive seasons. To explain adoption sequences, the approach requires assumptions about systematic changes in risk parameters or underlying preferences.

Ryan and Subrahmanyam found that the additional profits earned per unit of expenditures on a practice requiring "minimal change" was at least as great as the profit rate associated with the more complex, expensive packages. They argue that sequential adoption of components is a rational choice for a farmer with limited cash. As cash is accumulated from producing with the parts of the package, farmers will add components until the entire package has been adopted. This explanation suggests that a constrained farmer will adopt the choice which has the highest profit per dollar of expense, while an unconstrained farmer will adopt the choice which has highest profits.

Thus, a credit constraint can explain why a farmer may adopt a component even when the package is more profitable. However, the explanation exhibits two weaknesses. First, the model predicts that within a region, all credit-constrained farmers facing the same set of relative prices would adopt technological components in the same sequence. Second, when the input package displays "profit rate complementarity" (as defined by Ryan and Subrahmanyam, when the profit per dollar of expense is higher for the package than for any component), a credit constraint cannot explain why the farmer chooses

part rather than all of the package.¹ When profit rate complementarity exists, the credit-constrained farmer will adopt the whole package but perhaps on a smaller part of the land. If the additional assumption is added that the farmer must purchase the package in large, indivisible units, the credit constraint model can predict adoption of a part even when the package has the highest profit to cost ratio. By contrast, in the Mexican, Pakistani, and Zimbabwean cases cited above, the package inputs were described as highly divisible, variable inputs.

A Bayesian Approach to Technology Adoption

None of the above hypotheses is a consistent explanation of observed sequential adoption behavior. Cummings's observation that farmers experiment with recommendations suggests a learning model as a possible explanation. Many economists have used Bayesian approaches to explain aspects of technology adoption (O'Mara; Jensen; Hiebert; Stoneman; Feder; various papers by Lindner and others; Grossman, Kihlstrom, and Mirman).

Of special interest is Cyert and deGroot's model of financial decisions of a firm. In the model, the firm is considering investment in a large, fixed asset, and personnel may understate certain perceived risks associated with the project in order to obtain funding. Managers may choose to divide the investment into parts or modules that enable them to invest sequentially according to revised profitability expectations. Their decision is complicated by the fact that higher investment levels may provide more accurate information while increasing the investment risks. Cyert and DeGroot also show that it may be optimal for managers to invest partially (in some modules) if they can invest sequentially, even when the optimal investment would be zero in nonsequential conditions.

The notion that a choice will generate not only income but also information forms the basis of a behavioral model in which sequential adoption of new technology is a rational response. A general form of the model is presented in this section, followed by a simplified version illustrating the main points in the next section.

A producer will be in business T periods. Each

period, the producer chooses an action d_t from the decision set D_t . The choice each period generates stochastic income $y_t(d_t, \epsilon_t)$, where ϵ_t is the realization of a random variable ϵ with probability density function, f . The probability density function is not known with certainty but is known up to a vector θ . The decision maker has beliefs about what values θ can take; these beliefs are characterized by a probability density function $g_t(\theta; I_t)$ which is a function of the information possessed by the decision maker in period t (I_t). The information about θ includes the decision maker's past observations of d and y , as well as any other information i available at time t : $I_t = I_t(d_{\tau < t}, y_{\tau < t}, i_t)$. The probability distribution g_t is derived by a Bayesian update of the prior distribution, g_{t-1} , using the information set I_t .

The decision maker maximizes the present value of subjective expected utility (using discounting factor ϕ):

$$\text{Max}_{d_0, \dots, d_T} \sum_{t=0}^T \phi^t EU_t,$$

where $EU_t = \iint U(y_t(d_t, \epsilon))f(\epsilon, \theta)g(\theta, I_t) d\epsilon d\theta$. The problem is solved by backward induction, and the solution is the sequence $\{d_t^*\}$, for $t = 0 \dots T$.

The decision in any period is

$$\text{Max}_{d_0} E_t U_t(y_t(d_t)) + \sum_{i=t+1}^T E_t \text{Max}_{d_1, \dots, d_T} E_i U_i(I_i(d_i)),$$

where E_t denotes the expectation in period t and U_i denotes the utility in period i . This representation decomposes the effect of each period's choice into a direct effect influencing income in that period y_t , and the indirect effect of influencing the information available in future periods. The "full information optimum" is the choice d_t^{fi} which maximizes $\int U(y_t(d_t, \epsilon))f(\epsilon, \theta^{\text{fi}}) d\epsilon$, where θ^{fi} is the true (full information) value of θ . If full information becomes available at period \bar{t} , then for all $t \geq \bar{t}$, $g(\theta^{\text{fi}}) = 1$, $g(\theta \neq \theta^{\text{fi}}) = 0$, the informational content of any choice d_t is zero and $d_t^* = d_t^{\text{fi}}$. However, for $t < \bar{t}$, d_t^* will differ from d_t^{fi} whenever the sacrifice in expected utility in the current period is more than compensated by higher expected utility in future periods obtained with the better information.

This model is intentionally a very general model of producer decisions in the presence of risk and incomplete information. Sequential adoption of parts of a new technology is only one part of the decision-making process. Tech-

¹ Profit rate complementarity implies profit complementarity as defined by Feder, when profits of the package exceed profits of any component.

nology adoption consists of three distinct decisions within one optimization problem: a discrete choice of technological package, a continuous choice of land allocation between the new and traditional technology, and a choice of input levels. In the context of the above model, the choice set D_t may consist of elements which specify the components of new technology, the extent of new versus traditional agricultural techniques, and the intensity of adoption. For example, one element of D_t might be "use traditional growing techniques on two hectares, a new seed variety at the recommended seeding rate on one hectare, and the new seed variety at 80% of the recommended seeding rate on two hectares.

This paper focuses on the first of these three decisions: which components of the new technology to adopt. It is clear how this model offers an explanation for observed sequential adoption behavior. If the package is the combination of components which performs best among all possible combinations of new techniques, d_t^f will include the recommended package (perhaps along with traditional growing practices).

However, in periods $t \leq \bar{t}$, if the information gleaned from adopting a part of the package is a better bargain than information gleaned from adopting the package as a whole, d_t^f may include some component or components of the package, rather than the entire package. Certainly, if farmers can learn about the whole package by buying only a part of it, there will be an information incentive to buy the part rather than the whole.²

This model is consistent with all empirically observed facts. For the most part, individuals will move from the part to the whole package as they gain better information about the package. Occasionally, individuals will learn that the package is not appropriate for them. Normal adoption patterns will differ from region to region because of differences in direct effects and informational effects. Adoption patterns will differ from individual to individual within a region because of differences in the Bayesian priors of individuals in period zero. All of these points

are illustrated in the simple illustration presented below.

A Simple Illustration

The general model of learning relies on two basic conditions: the farmer is uncertain about the net returns associated with adopting a new technological package, and the uncertainty may not be resolved by adopting the package. The farmer operates for two production periods ($T = 2$), and decides in each period whether to adopt a new seed variety, a commercial fertilizer, or both. Thus the decision set contains three elements: $D_t = \{S, F, B\}$, where S represents the adoption of seed, F the adoption of fertilizer, and B the adoption of both. The decision is influenced by information that an extension agent provides about the effectiveness of the techniques. The extension agent claims that the additional costs and returns associated with adoption of the new techniques are those shown in table 1. R_S, R_F, R_B are expected revenues from the possible choices, C_S, C_F, C_B are costs of the choices, α, β, γ are probabilities, and s, f , and b are constants.

The farmer is not certain that these figures correctly indicate how new techniques will perform on his/her fields.³ If the extension agent's report is not valid for his/her farm, the farmer believes that additional costs and returns are shown in table 2.

Both of the basic conditions of the model are met. First, the farmer is uncertain about which of the two tables is appropriate to his/her farm. Second, because the "low revenue" when the report is valid is identical to the "high revenue" when the report is not valid, farming experience will not always eliminate uncertainty. If this ambiguous level of revenues is observed, the farmer cannot distinguish whether it expresses a high revenue and an invalid report or low revenue and a valid report.

Superscripts 1 and 2 refer to periods, and subscripts S, F , and B denote the adoption of seed only, fertilizer only, or both seed and fertilizer. Expected profits earned in a period by the farmer, conditional on a valid extension report, are de-

² The decision of how extensively to apply the new (experimental) technique will be influenced by the farmer's subjective judgment about its ultimate applicability. If the farmer is doubtful about whether the new technology is an improvement on traditional practices, he/she may try the new technique on only a small portion of land. This will be especially true if the farmer can learn as much from trying the new technique on a small versus a big plot.

³ A farmer's experience may differ from extension claims because of differences in objectives, agroecological conditions, or related technology (Herd). Research institutions work with "groups of farmers for whom the frequency distributions of expected yield increments will not be very compact due to agroclimatic variability" (Perrin and Winklemann, p. 892).

Table 1. Additional per Hectare Costs and Returns of New Techniques If Extension Report Is Valid

Technique	Low Revenue	Probability of Low Revenue	High Revenue	Probability of High Revenue	Expected Revenue	Cost
Seed	$R_s - s/(1 - \alpha)$	$(1 - \alpha)$	$R_s + s/\alpha$	α	R_s	C_s
Fertilizer	$R_f - f/(1 - \beta)$	$(1 - \beta)$	$R_f + f/\beta$	β	R_f	C_f
Both	$R_b - b/(1 - \gamma)$	$(1 - \gamma)$	$R_b + b/\gamma$	γ	R_b	C_b

noted by π (e.g., $\pi_B = R_B - C_B$). Expected profits conditional on an invalid extension report are denoted by μ (e.g., $\mu_B = (1 - \gamma)R_B - b - C_B$). The farmer believes that the report is valid with probability δ . The farmer's subjective unconditional expected profit is denoted by ψ (e.g., $\psi_B = \delta\pi_B + (1 - \delta)\mu_B$).

In this example, farmers are assumed risk neutral. In addition, certain relationships between inputs are also assumed in constructing the model, and are contained in the expressions:

$$(1) \quad \pi_B > \pi_S > 0, \quad \pi_B > \pi_F > 0;$$

$$(2) \quad \mu_S < \mu_B < 0, \quad \mu_F < \mu_B < 0.$$

The first assumption formalizes the notion that the components of the technological package are profit complementary and implies that adopting either part of the package is preferred to the traditional technique when the extension report is valid.⁴ The second assumption states that the components are complementary even when the extension report is invalid and implies that under those conditions the (risk-neutral) farmer prefers current techniques. The two assumptions, combined, indicate that the farmer's unconditional expected profits are higher with the package than with any part, or

$$(3) \quad \psi_B > \psi_s \text{ and } \psi_B > \psi_F.$$

⁴ Unlike the credit constraint explanation of sequential adoption, we make no assumptions about profit complementarity: $\Pi_B/C_B > \Pi_s/C_s > \Pi_F/C_F$ is permitted, but not required, by this formulation.

For ease of exposition, δ is defined to range in magnitude close enough to 1 that $\psi_B > 0$. The inequality can be reversed without affecting the fundamental arguments generated by the model.

The farmer operates for two periods, using experience gained in the first period to guide his/her choice in the second by updating, according to Bayes's rule, the subjective judgment about the value of δ . When only seed is adopted in the first period, the farmer faces one of three possible outcomes. An outcome of $(R_s + s/\alpha)$ proves that the extension report is valid, and the farmer updates the first-period subjective probability (δ^1) to $[\delta^2 = 1]$ in the second period. An outcome of zero revenue confirms that the extension agent is wrong, and the farmer's updated subjective probability is $[\delta^2 = 0]$. In either of the two outcomes, the first-period experiment eliminates uncertainty in the second period.

By contrast, an outcome of $R_s - [s/(1 - \alpha)]$ is ambiguous: the farmer cannot determine whether the outcome occurred because the extension report is invalid, and the technology performed well; or because the report is valid, but the technology performed poorly. When first-period choice results in this ambiguous level of revenue, the farmer's updated subjective probability is $\delta^2 = [\delta^1(1 - \alpha)]/[(1 - \alpha)(\delta^1 + 1 - 0\delta^1)] = \delta^1$. In other words, the updated subjective probability is the same as the prior. Analogous alternative outcomes exist for a technology choice of fertilizer only or the package of both seed and fertilizer.

As the problem is formulated, the experience

Table 2. Additional per Hectare Costs and Returns of New Techniques If Extension Report Is Not Valid

Technique	Low Revenue	Probability of Low Revenue	High Revenue	Probability of High Revenue	Expected Revenue	Cost
Seed	0	α	$R_s - s/(1 - \alpha)$	$(1 - \alpha)$	$(1 - \alpha)R_s - s$	C_s
Fertilizer	0	β	$R_f - f/(1 - \beta)$	$(1 - \beta)$	$(1 - \beta)R_f - f$	C_f
Both	0	γ	$R_b - b/(1 - \gamma)$	$(1 - \gamma)$	$(1 - \gamma)R_b - b$	C_b

of the first period will completely eliminate uncertainty in two cases: when revenue greater than R_i (expected revenue when alternative i is chosen and the extension report is valid) is observed, the farmer knows that the extension report must be valid; when revenue of zero is observed, the farmer knows that the extension report must not be valid. The probability of attaining certainty through the first year's experience is α for farmers who adopt seed only, β for farmers who adopt fertilizer only, and γ for farmers who adopt both seed and fertilizer. The probabilities of observing the ambiguous, middle revenue are, respectively, $(1 - \alpha)$, $(1 - \beta)$, and $(1 - \gamma)$.

The crux of this example is that observing middle revenue does not inform the farmer about the validity of the extension report. If that ambiguous revenue occurs, the possibility remains that the second-period choice will be "wrong." In the second period, the farmer may choose to adopt the package although the extension report is invalid or, conversely, may not adopt even though the extension report is valid. The farmer's decision about production techniques in period one is, therefore, influenced not only by the profits expected in that period but also by the informational content of the choice. Better information acquired during the first period reduces the chance of making the wrong choice in the second period. The farmer recognizes that the second-period decision will be affected by the technology choice in the first period.

The risk-neutral farmer's objective is to choose the production technique in each period from a set that includes the traditional technique, seed only, fertilizer only, and the seed-fertilizer package, knowing that the decision in the second period is affected by the first-period choice. The choice is conditional on the farmer's subjective judgment about the probability that the extension report is valid. Some choices may be more informative about the package than others. Viewed from the first period, the farmer maximizes expected profits over the two periods and over the set of possible outcomes.

The farmer's problem is solved by backward induction. First, the optimal choice in the second period is identified conditional on the value of δ^2 . The parameter δ^2 takes one of three values. When $\delta^2 = 0$, the farmer is certain that the extension report is invalid; he/she chooses the traditional farming techniques in the second period, and earns zero additional profit per hectare. When $\delta^2 = 1$, the farmer is certain that the extension report is valid; he/she chooses the seed-

fertilizer package and earns additional expected profit of $\pi_B = R_B - C_B$. When $\delta^2 = \delta^1$, the farmer remains uncertain about the validity of the extension report, but he/she chooses the seed-fertilizer package because, by assumption, ψ_B is positive and larger than either ψ_S or ψ_F .

The expectation the farmer forms in the first period, E^1 , can then be expressed in terms of each of the possible first period choices. If the farmer chooses S in the first period, total expected profit viewed from period 1 is

$$\begin{aligned} E_S^1 &= \psi_S + \delta^1 \alpha \pi_B + \delta^1 (1 - \alpha) \pi_B \\ &\quad + (1 - \delta^1) \alpha \cdot 0 + (1 - \delta^1) (1 - \alpha) [\mu_B], \\ &= \psi_S + \delta^1 \pi_B + (1 - \delta^1) (1 - \alpha) [\mu_B]. \end{aligned}$$

If the farmer chooses F or B in the first period, expected profits are

$$\begin{aligned} E_F^1 &= \psi_F + \delta^1 \pi_B + (1 - \delta^1) (1 - \beta) [\mu_B] \\ &\quad \text{(for fertilizer only)} \\ E_B^1 &= \psi_B + \delta^1 \pi_B + (1 - \delta^1) (1 - \gamma) [\mu_B] \\ &\quad \text{(for the package).} \end{aligned}$$

The first term in each expression is expected return in the first period from using the indicated technique. The remaining terms show expected return in the second period conditional on the first-period choice.

The case of the farmer's decision to adopt only seed in the first period illustrates the problem. If the extension report is valid (with subjective probability δ^1), the farmer obtains revenue higher than R_i with a probability of α and observes the ambiguous revenue level with probability $(1 - \alpha)$. With either outcome, the farmer adopts the seed-fertilizer package in the second period, and earns expected additional profits of π_B . If the extension report is invalid [probability $(1 - \delta^1)$], the farmer earns zero revenue with probability α and the ambiguous revenue level with probability $(1 - \alpha)$ in the first period. When zero revenue occurs, the farmer reverts to the traditional technique in the second period and earns additional profits of zero. Observing middle revenue leads the farmer to adopt (mistakenly) the seed-fertilizer package in period 2 and to earn (negative) additional profits of μ_B .

Behavioral Implications

The above example demonstrates how the Bayesian model explains, without imposing institutional constraints or assuming risk aversion, why farmers may adopt components of a tech-

nological package sequentially even when the profitability of adopting the package is greater than that associated with the use of a component. In the first period, the farmer uses the technique which yields the highest expected return. A choice of seed only implies that $E_S^1 > E_B^1$, or

$$E_S^1 - E_B^1 = [\psi_S - \psi_B] + (1 - \delta^1)\mu_B[\gamma - \alpha] > 0.$$

The difference in the expectations consists of two terms.

The first term $[\psi_S - \psi_B]$ represents the difference in expected period one profits when the farmer chooses seed rather than the package. The term is negative by assumption; we are interested here in the case where the package is the "best" choice in the long run. The second term expresses the difference in expected costs caused by a mistaken choice in the second period. The probability associated with the mistake is $(1 - \alpha) - (1 - \gamma) = (\gamma - \alpha)$, and μ_B , which is negative by assumption, is the cost of choosing to adopt the package in the second period when the "correct" (highest expected profit) choice is the traditional technique. As the problem is depicted, a necessary condition for adoption of only seed in the first period is $\gamma < \alpha$, or that probability of earning high revenue with the package, when the extension report is valid, is less than the probability of earning high revenue with seed only. Because γ and α also represent the probability of eliminating uncertainty under each choice, the condition $\gamma < \alpha$ means that the package is less informative than seed only. For a farmer to sequentially adopt seed first and then the seed-fertilizer package, the adoption of seed only must be more informative about the validity of the agent's report than the adoption of the package. Analogous results hold for adoption of fertilizer only.

The model reveals some interesting insights into the adoption patterns described in empirical work. A δ^1 sufficiently close to one, which denotes high confidence in the success of the package, ensures that sequential adoption will not occur. If the extension agent is right, and when farmers have other sources of information such as neighboring farmers, the model suggests that early adopters may adopt parts of the package sequentially, while later adopters, whose δ^1 's have been raised by the positive experience of neighbors, adopt the whole package.

The model also demonstrates why farmers in one region tend to adopt a part of the package first, while farmers in another region tend to adopt

another. In circumstances where $E_S^1 > E_B^1$ and $E_F^1 > E_B^1$, the farmer will adopt seed first if $E_S^1 - E_F^1 > 0$ and fertilizer first if $E_S^1 - E_F^1 < 0$. The order of adoption is determined by the sign of

$$E_S^1 - E_F^1 = [\psi_S - \psi_F] + (1 - \delta^1)\mu_B(\beta - \alpha),$$

and will vary between two regions if the corresponding ψ_S , ψ_F , μ_B , α , and β differ. Differences in the general pattern of adoption are explained by differences in the economic or agroclimatic environment that affect the profit distribution associated with each choice of technique.

The order of adoption, or the sign of $E_S^1 - E_F^1$, is also affected by δ^1 , the farmer's perceived probability that the extension report is accurate. If $\psi_S > \psi_F$ and $\beta > \alpha$, seed alone generates higher expected returns in the first period, but fertilizer alone provides more information. In this case, a farmer with a low δ^1 may choose fertilizer in the first period, while another farmer, identical to the first except for δ^1 , chooses seed. Farmers with similar characteristics but differing prior judgments about the reliability of the extension report may choose to adopt different components. Thus, the model can explain why farmers facing the same economic and ergonomic conditions choose different adoption sequences.

Finally, the model indicates that farmers may choose to adopt a particular technique because they are willing to pay for information to avoid making a mistake in subsequent periods. For example, when ψ_B is negative, the farmer's judgment is that expected profits in a period will be maximized by using traditional production techniques. The farmer may nevertheless adopt a part or all of the package in the first period if the expectation, E^1 , is positive for that choice. The farmer is willing to pay for information in the first period to avoid making a mistake in the second period.

The analytical results of the model are not grounded in assumptions about the degree of input complementarity, and even when inputs demonstrate profit-rate complementarity, all of the cases described can be realized. As formulated, the results do not depend on capital or credit constraints that curtail input purchases or on risk-averse behavior.

Concluding Comments

Sequential or stepwise adoption of parts of a technological package has been observed in a

variety of settings, but the reasons behind it are only partially understood. The model presented in this paper is consistent with the behavior described by informed observers; it presents a plausible explanation for sequential adoption even under the assumptions of risk neutrality and unconstrained expenditures. According to this explanation, a farmer is uncertain about whether the new production techniques will be successful on his/her farm, and experimentation with techniques will not always resolve this uncertainty. Under some circumstances, the farmer may learn more about the innovation as a whole by adopting only a part of the recommended package. The model can explain why patterns of sequential adoption differ among agroecological regions, as well as among neighboring farmers operating under similar conditions. Unlike the credit constraint explanation described earlier, the Bayesian approach explains sequential adoption even when the profit rate from adopting the package is greater than with any component.

The model supports the perspective that sequential adoption may be a rational choice for imperfectly informed farmers and need not be viewed as evidence of institutional shortcomings or farmer irrationality. If sequential adoption is a response to uncertainty on the part of farmers, the appropriate policy question is whether sequential adoption is the least-cost way of providing information to farmers.

[Received November 1989; final revision received October 1990.]

References

- Byerlee, D., and E. Hesse de Polanco. "Farmers' Stepwise Adoption of Technological Packages: Evidence from the Mexican Altiplano." *Amer. J. Agr. Econ.* 68(1986):519-27.
- Centro Internacional de Mejoramiento de Maiz y Trigo (CIMMYT). Draft of "1987-1988 World Wheat Facts and Trends: the Wheat Revolution in the Third World: Past, Present, and Future." El Batán, Mexico, 1988.
- Cummings, Ralph W., Jr. "The Puebla Project." Paper prepared for meeting on Social Science Research in Rural Development, Rockefeller Foundation, 29-30 April 1975 (cited in Ryan and Subrahmanyam).
- Cyert, Richard M., and Morris H. DeGroot. "Sequential Investment Decisions." *Bayesian Analysis and Uncertainty in Economic Theory*, chap. 11. Totowa NJ: Rowman and Littlefield, 1987.
- Feder, G. "Adoption of Interrelated Agricultural Innovations: Complementarity and the Impacts of Risk, Scale, and Credit." *Amer. J. Agr. Econ.* 64(1982):94-101.
- Feder, G., R. Just, and D. Zilberman. "Adoption of Agricultural Innovations in Developing Countries: A Survey." *Econ. Develop. and Cultur. Change* 33 (1985):255-97.
- Grossman, Sanford J., Richard E. Kihlstrom, and Leonard J. Mirman. "A Bayesian Approach to the Production of Information and Learning by Doing." *Rev. Econ. Stud.* 43(1976):533-47.
- Herd, R. W. "Increasing Crop Yields in Developing Countries: Sense and Nonsense." Paper presented at AAEA annual meeting, Knoxville TN, 30 July-3 Aug. 1988.
- Heisey, P. W., et al. "Diagnosing Research Priorities for Higher Altitude Maize-Based Farming Systems in Swat." Draft report, Pakistan Agricultural Research Council/CIMMYT, Islamabad, 1988.
- Hiebert, D. "Risk, Learning, and the Adoption of Fertilizer Responsive Seed Varieties." *Amer. J. Agr. Econ.* 56(1974):764-68.
- Husain, Tariq. "Innovation Adoption and Choice of Technique in a Mountain Farming System: A Study of New Wheat Varieties in the Gilgit District of Pakistan." Ph.D. thesis, University of Chicago, March 1987.
- Jensen, Richard. "Adoption and Diffusion of an Innovation of Uncertain Profitability." *J. Econ. Theory* 27 (1982):182-93.
- Lele, U., and A. Goldsmith. "The Development of National Agricultural Research Capacity: India's Experience with the Rockefeller Foundation and Its Significance for Africa." *Econ. Develop. and Cultur. Change* 37(1989):305-43.
- Lindner, Robert K. "Farm Size and the Time Lag to Adoption of a Scale Neutral Innovation." Mimeographed. Adelaide, Australia: University of Adelaide, 1980.
- Lindner, Robert K., and A. J. Fischer. "Risk Aversion, Information Quality, and the Innovation Adoption Time Lag." Mimeographed. Adelaide, Australia: University of Adelaide, 1981.
- Lindner, Robert K., A. J. Fischer, and P. Pardey. "The Time to Adoption." Mimeographed. Adelaide, Australia: University of Adelaide, 1979.
- Lowdermilk, M. K. "Diffusion of Dwarf Wheat Production Technology in Pakistan's Punjab." Summary report condensed from Ph.D. thesis, Cornell University, 1972.
- Mann, C. K. "Packages of Practices: A Step at a Time with Clusters." Middle East Technical Institute: *Studies in Development (Gelisme Dergisi)*, 21(1978):73-82.
- Moseman, Albert. *Building Agricultural Research Systems in the Developing Nations*. New York: Agricultural Development Council, 1970.
- O'Mara, G. "The Microeconomics of Technique Adoption by Smallholding Mexican Farmers." *The Book of Chac: Programming Studies for Mexican Agriculture*, ed. Roger D. Norton and Leopoldo Solis M. Baltimore MD: Johns Hopkins University Press and the World Bank, 1983.
- Perrin, R., and D. Winkelmann. "Impediments to Technical Progress on Small versus Large Farms." *Amer. J. Agr. Econ.* 58(1976):888-94.

On Testing the Structure of Risk Preferences in Agricultural Supply Analysis

Rulon D. Pope and Richard E. Just

Risk preferences broadly affect many economic decisions when markets are incomplete. Common representations of risk preferences are constant absolute, relative, and partial relative risk aversion. Each of these preference classes has distinct impacts on choice. An econometric test for distinguishing the class of preferences is proposed and implemented for potato supply response in Idaho. The data reject constant absolute and partial relative risk aversion and are congruent with constant relative risk aversion.

Key words: risk, risk preferences, supply response.

The model of the price-uncertain, competitive, single-product firm has received much attention in a wide variety of theoretical analyses (e.g., Sandmo, Katz, Briys and Eeckhoudt, Ishii, Hey, Chavas and Pope). A rather complete set of theoretical restrictions has emerged which depends crucially on risk preferences. Three classes emerge as important benchmarks: constant absolute risk aversion (CARA), constant relative risk aversion (CRRA), and constant partial relative risk aversion (CPRRA). These three classes are important not only for null hypotheses from which to accept (or reject) rising or falling risk aversion but are important in their own right. For example, these three classes imply, respectively, (a) changes in the location of initial wealth do not alter decisions, (b) changes in the scale of wealth do not affect decisions, and (c) changes in profit or the scale of the bet do not alter decisions. Each of these implications has important ramifications. For example, CARA is crucial to exact welfare analysis (Pope, Chavas, and Just), while CRRA implies that a proportional wealth tax or wealth sharing does not alter decisions. For CPRRA, the scale of, or the existence of, profit sharing or tax would not affect decisions (Pope).

In addition to this conceptual work, a growing body of empirical work has found evidence of risk aversion and risk response. This includes both firm-level and aggregate analysis (compare

Binswanger, Just, Holland). In both approaches, important information is obtained. Firm-level analysis suffers from less aggregation bias and filters and controls extraneous factors more directly. However, aggregate analysis generally gives the kinds of measures that are needed for policy analysis. The econometric evidence in Hansen and Singleton, Friend and Blume, Siegel and Hoban, Wolf and Pohlman, and Szpiro finds support for risk-averse behavior; but whether risk aversion is constant, decreasing, or increasing remains ambiguous. Decreasing absolute risk aversion has emerged as a "stylized" fact or belief, but the nature of relative and partial relative risk aversion is in dispute.

This paper focuses on the econometric approach to risk response behavior for both the individual and aggregate levels and both the single and multiproduct cases. Simple statistical tests for CARA, CRRA, and CPRRA are discussed and applied. These tests use only reduced-form econometric models of the type commonly employed in agricultural economics research. They are implemented only within a single equation.¹ These tests are applied to a potato acreage response model for the state of Idaho. Strong evidence is found against the hypotheses of risk neutrality, CARA, and CPRRA. The data lends credence to the hypothesis of CRRA.

Rulon D. Pope is a professor of economics, Brigham Young University; Richard E. Just is a professor of agricultural and resource economics, University of Maryland.

¹ This is important because integrability of decision functions (rationalizing expected utility from multiple choices) in a production system is only attained in very specific risk models (e.g., Just and Pope, Antle).

The Firm's Decisions

The model used here is the netput model of the competitive firm (McFadden). Profit, π , is $p'X$, where p is a vector of random prices and X_i is a supply ($X_i > 0$) or derived-demand level ($X_i < 0$), $i = 1, \dots, N$. Initial wealth is assumed nonrandom and is denoted by W_0 . Hence, final wealth, W , is $W = W_0 + \pi$. Decisions are made *ex ante* by maximizing expected utility of wealth, $E[U(W)]$, subject to the feasibility condition (production possibility set and/or capital constraint) $T(X) = 0$, where U is utility; $E[U(W)]$ and T are assumed smooth in X and relevant moments of the random price vector, p .

Given these assumptions, the optimal choices are characterized by

$$(1) \quad \begin{aligned} E[U_w(W)p] + \lambda T_X(X) &= 0, \\ T(X) &= 0, \end{aligned}$$

where subscripts denote derivatives and λ is a Lagrangian multiplier. Solution of (1) yields optimal decisions X^* .

In empirical work, the distribution of p is assumed to be summarized parametrically by $\{M_1, \dots, M_s\}$, where each M_s is a vector representing all moments of order s .² In this case, direct expected utility can be expressed as $E[U(W)] = V(X, W_0, M_1, \dots, M_s)$ with optimal decisions.

$$(2) \quad X^*(W_0, M_1, \dots, M_s).$$

The essential nature of the problem here is that W is linear in p (but not X). Thus, the basic capital budgeting or portfolio problem is a special case.³

The three risk aversion measures explored here are

$$\frac{-U''(w)}{U'(w)} \quad \text{absolute risk aversion}$$

$$\frac{-U''(w)w}{U'(w)} \quad \text{relative risk aversion}$$

$$\frac{-U''(w)\pi}{U'(w)} \quad \text{partial relative risk aversion}$$

The first two measures are found in Arrow and Pratt, and the latter is from Menezes and Hansen.

A number of authors have applied expected utility models of the firm to profit instead of wealth (e.g., Sandmo, Binswanger, Antle). In such cases relative and partial risk aversion are identical. Katz has argued that any measure using only profit as the argument of utility (such as partial relative risk aversion) is inappropriate since profit can be negative. Hey refutes Katz but shows that the possibility of negative profit merely imposes restrictions on U . For example, constant absolute risk aversion yields risk-prefering behavior in the region of negative profit (Hey).

Testing for Specific Forms of Risk Aversion

As Pratt, Sandmo, and others have observed, changes in W_0 leave X^* unchanged under CARA. This provides a straightforward test for CARA. Turning to CRRA, a simple test can be developed by examination of (1). From Pratt (p. 134), if wealth is scaled by γ , $U(\gamma W)$ and $U(W)$ represent the same ordering under CRRA. This leads to the class of Cobb-Douglas utility functions in wealth. Therefore, if wealth is scaled by γ , the first-order conditions in (1) are effectively unaltered as is X^* . A similar argument can be made for CPRRA (Menezes and Hansen). If π is scaled by γ under CPRRA, X^* is unaltered. This leads to the class of Cobb-Douglas utility functions in profit.

These results hold regardless of the form of W . However, writing W as $W_0 + p'X$ allows simple and specific behavioral propositions which can be implemented easily. Scaling terminal wealth by γ is equivalent to scaling both initial wealth, W_0 , and prices, p , by γ . Furthermore, scaling p by γ is equivalent to scaling all moments M_s by γ^s .⁴

These results imply

² The moments are assumed to exist. Most empirical work includes only M_1 and M_2 , for example, where the class of stochastic distributions is restricted to normality (see the empirical references cited above). The methodology here is more generally restricted to a class of stochastic distributions that are completely characterized by a finite set of moments.

³ For example, in one special case the individual or firm solves $\max E[U(W)]$ subject to $\sum_{i=1}^N X_i \leq W_0$, where p_i is interpreted as the return on one invested dollar in activity i . In a common agricultural case, the farmer solves $\max E[U(W)]$ subject to $\sum_{i=1}^N X_i \leq \bar{A}$, where \bar{A} is available acreage, X_i is the acreage allocated to crop i , and p_i is the net returns per acre from farming crop i .

⁴ Where an element of M_s is given by

$$\begin{aligned} M_s^k(p) &= E(p_1 - E(p_1))^{\alpha_1} (p_2 - E(p_2))^{\alpha_2} \dots (p_N - E(p_N))^{\alpha_N}, \\ \sum_{i=1}^N \alpha_i &= s, \text{ scaling } p = (p_1, \dots, p_N)' \text{ by } \gamma \text{ obtains } M_s^k(\gamma p) \\ &= \gamma^s M_s^k(p). \end{aligned}$$

$$(3) \quad X^*(\gamma W_0, \gamma M_1, \gamma^2 M_2, \dots, \gamma^s M_s) \\ = X^*(W_0, M_1, M_2, \dots, M_s).$$

That is, under CRRA decisions in (2) are almost homogenous of degree one in initial wealth and the means of p , of degree two in the second moment of p , and degree s in the s th moment of p .⁵ As in common tests for homogeneity, a Euler equation corresponding to (3) is useful for testing CRRA.

PROPOSITION 1. *Under CRRA, decisions in (2) follow the vector equation:*

$$(4) \quad \left(\frac{\partial X^*}{\partial W_0} \right) W_0 + \sum_{s=1}^s \left(\frac{\partial X^*}{\partial M_s} \right) M_s = 0, \text{ or} \\ \left(\frac{\partial \ln X^*}{\partial \ln W_0} \right) + \sum_{s=1}^s \left(\frac{\partial \ln X^*}{\partial \ln M_s} \right) s = 0,$$

where appropriate.

Proof: Differentiate (3) with respect to γ and evaluate at $\gamma = 1$ (Pope).

Thus, for example, if one estimates (2) in linear logarithmic form, the weighted sum of the elasticities in (4) must equal zero. The weights are the order of the moments. Equation (4) can be imposed locally by imposing (4) at a point $(\bar{W}_0, \bar{M}_1, \dots, \bar{M}_s)$.

The analysis for CPRRA is similar. Scaling π by γ is equivalent to scaling all moments M_s by γ^s , which implies

$$(5) \quad X^*(W_0, \gamma M_1, \dots, \gamma^s M_s) \\ = X^*(W_0, M_1, \dots, M_s).$$

Thus, under CPRRA, decisions in (2) are almost homogenous of degree one in the means of p , of degree two in the second moment of p , and degree s in the s th moment of p .

PROPOSITION 2. *Under CPRRA, decisions in (2) follow the vector Euler equation*

$$(6) \quad \sum_{s=1}^s \left(\frac{\partial X^*}{\partial M_s} \right) M_s = 0.$$

Proof: Similar to proposition 1.

Again (6) can be imposed locally for local tests of CPRRA. Finally, a well-known result is stated.

PROPOSITION 3. *Under CARA, initial wealth does not alter decisions, or*

$$(7) \quad \frac{\partial X^*}{\partial W_0} = 0.$$

A simple example will illustrate how propositions 1, 2, and 3 can be applied in testing for CARA, CRRA, and CPRRA. Suppose that p_1 is random and $X_1^* > 0$ with p_2, \dots, p_N nonrandom and $X_2^*, \dots, X_N^* < 0$. This is the Sandmo model. Assuming p_1 is normally distributed, decisions are completely characterized by the mean of p_1 , $E(p_1) = \mu$; σ^2 , the variance of p_1 ; p_2, \dots, p_N ; and W_0 . Under CARA, expected utility is maximized by equivalently maximizing $\mu X_1 + \sum p_j X_j - (r/2) X_1^2 \sigma^2$ subject to $T(X) = 0$ (Freund). Thus, $\partial X^* / \partial W_0 = 0$ since initial wealth does not enter the decision. Further, because the objective function in this particular case is linearly homogenous in μ , p_2, \dots, p_N and σ^2 , the following three tests under normality are appropriate:

$$(8) \quad \frac{\partial W^*}{\partial W_0} = 0, \frac{\partial X^*}{\partial \mu} \mu + \sum_{j=2}^N \left(\frac{\partial X^*}{\partial p_j} \right) p_j \\ + \left(\frac{\partial X^*}{\partial \sigma^2} \right) \sigma^2 = 0 \text{ for CRRA}$$

$$(9) \quad \frac{\partial W^*}{\partial W_0} W_0 + \left(\frac{\partial X^*}{\partial \mu} \right) \mu + \sum_{j=2}^N \left(\frac{\partial X^*}{\partial p_j} \right) p_j \\ + 2 \left(\frac{\partial X^*}{\partial \sigma^2} \right) \sigma^2 = 0 \text{ for CARA}$$

$$(10) \quad \left(\frac{\partial X^*}{\partial \mu} \right) \mu + \sum_{j=2}^N \left(\frac{\partial X^*}{\partial p_j} \right) p_j \\ + 2 \left(\frac{\partial X^*}{\partial \sigma^2} \right) \sigma^2 = 0 \text{ for CPRRA.}$$

The first restriction in (8) always occurs under CARA, whereas the second restriction is specific to CARA under normality. It is interesting to contrast the Euler equations in (8)–(10) with those occurring under certainty or risk neutrality. Under certainty or risk neutrality, all wealth and variance effects are omitted. Thus, both certainty and risk neutrality are special cases of all three conditions. Finally, note that the variance term has weight one under CARA and two under CRRA and CPRRA.

An Application

The risk response models developed by Just and others are applied here to potato production in

⁵ Here we use the concept of almost homogeneity used, for example, by Lau (p. 173): A function $h(z_1, \dots, z_N)$ is said to be almost homogenous of degree c_1, \dots, c_N , and zero if $h(\gamma^{c_1} z_1, \dots, \gamma^{c_N} z_N) = \gamma h(z_1, \dots, z_N)$ where $\gamma \in R$. Here we consider γ to be positive and is thus positively almost homogenous.

Idaho to test for the structure of risk preferences in parsimonious reduced-form models such as (8)–(10). Potatoes have been of longstanding interest to economists; their highly erratic prices have made potatoes a prime candidate for the analysis of perverse demand behavior (Dwyer and Lindsay), dynamic supply analysis (Zusman), and risk response (Estes, Blakeslee, and Mittelhammer).

The model specification is

$$(11) \quad A_p = B_0 + B_1 W_0 + B_2 P_p^e + B_3 P_s^e + B_4 V_p^e + B_5 V_s^e + B_6 C^e + \epsilon,$$

where all variables are T vectors with $\epsilon \sim N(0, \sigma^2 I_T)$, i.e., ϵ_t is identically and independently distributed, A_p is potato acreage, the B 's are parameters to estimate; the subscripts p and s represent potatoes and sugar beets, respectively; and superscript e represents the subjectively formed moments of output prices, where P^e is mean price, V^e is variance of price, and C^e is covariance of prices.⁶ Letting t denote time, the first two moments are defined as in Just:⁷

$$(12) \quad P_{jt}^e = \theta \sum_{k=0}^K (1 - \theta)^k P_{j,t-k-1}; j = s, p,$$

$$(13) \quad V_{jt}^e = \theta \sum_{k=0}^k (1 - \theta)^k (P_{j,t-k-1} - P_{j,t-k-1}^e)^2;$$

$$j = s, p,$$

$$(14) \quad C_t^e = \theta \sum_{k=0}^k (1 - \theta)^k \{ [P_{j,t-k-1} - P_{j,t-k-1}^e] [P_{j',t-k-1} - P_{j',t-k-1}^e] \} j \neq j', j = s, p$$

and can be rationalized as a Bayesian learning process where θ is a parameter. Though the debate over rational versus adaptive expectations leads to important criticisms of the adaptive model, the choice of adaptive expectations is necessitated by the fact that market-level data are not analyzed. Further, there is no futures market for potatoes, and thus using spot prices to form expectations for the price of potatoes is reasonable.

The time-series data are Idaho state-level data extending from 1950 through 1987.⁸ The first

six observations are used as a history for the adaptive expectations framework. Thus, the model is estimated with thirty-two observations. The measure of wealth is the value of land and buildings minus associated debts. Total state potato acreage is the dependent variable, and all the independent variables in (11) are normalized by prices paid for agricultural inputs in production. All variables are then normalized by their respective means. This implies that parameter estimates in a linear regression are elasticities at the average level of the variables.

The results from nonlinear regression are presented in table 1. The adaptive parameter θ is estimated in the unit interval with a fairly high level of memory indicated in the expectation scheme. The coefficients, $B_0 - B_4$ and θ have the expected signs. The sign of B_5 is counterintuitive when the covariance term is included but is insignificant. It changes to a plausible sign when the covariance term is excluded. In general, based upon the signs and precisions of coefficient estimates, the model with covariance excluded seems better specified.

The sign of B_6 in the first column is indeterminate so a positive value or zero is not implausible. That is, an increase in covariance tends to reduce diversification and leads to more specialization in the crop with higher expected returns. Potatoes are apparently that crop. Generally t -ratios are large, lending strong evidence (99% confidence) that wealth and own expectations and variance are significant explanatory variables. The evidence is less strong for inclusion of cross-price effects.

Based upon the individual t -tests for $H_0: B_4, B_5, B_6 = 0$, it appears that the hypothesis of risk neutrality ($B_4 = B_5 = B_6 = 0$) is soundly rejected. Indeed, a likelihood ratio test of the joint hypothesis yields a statistic of 9.6. This exceeds the chi-square tabled value of one degree of freedom at the .01 significance level. Wealth is positively related to potato acreage ($B_1 > 0$) and is highly significant, *ceteris paribus*. Thus, CARA is rejected in favor of nonconstant absolute risk aversion. The results also suggest rejection of CPRRA. Thus, omission of wealth effects in empirical research may bias estimates of responses including risk effects. Because such wealth effects are routinely omitted in agricultural research, opportunities for exploring risk preferences and measuring properly risk response properly are lost.

Table 2 contains the test results of the three hypotheses similar to (8)–(10). As noted earlier, these results soundly reject the hypothesis of

⁶ Antle (1987) has stressed that inclusion of higher moments should be investigated. However, convergence with even two moments is difficult.

⁷ All estimations assumed a single θ parameter in (12)–(14). The use of a single θ parameter is supported by the work of Just.

⁸ For sources, see data Idaho Crop and Livestock Reporting Service and U.S. Department of Agriculture.

Table 1. Potato Supply Response—Idaho

	Coefficient	With Covariance	Without Covariance
Intercept	B_0	-.2350 (.4559)	.0192 (.2452)
Wealth	B_1	.6831 (.0774)	.6651 (.0802)
Price expectation—potatoes	B_2	2.6548 (.7327)	2.9378 (.6537)
Price expectation—sugar beets	B_3	-1.2209 (1.7610)	-2.0751 (.2314)
Variance—potatoes	B_4	-1.6801 (.61433)	-1.1971 (.330)
Variance—sugar beets	B_5	-.7385 (1.8857)	1.0103 (.2205)
Covariance	B_6	2.2339 (2.0709)	
Adaptive parameter	θ	.3258 (.1540)	.4205 (.2169)
SSR		.2310	.2438
R^2		.77	.77

Notes: Standard errors in parentheses are computed using the information matrix. The Hessian of the sum of squares is divided by $2\hat{\sigma}^2$. The negative of the inverse of the resulting matrix is the information matrix derived from the log likelihood function. Coefficient estimates are elasticities at sample means.

CARA at the 1% level. They also reject the hypothesis of CPRRA at the 1% level. Not only does wealth enter utility, but the Euler equation in (6) significantly raises the likelihood function. The hypothesis of CRRA cannot be rejected at this level. In the model with covariance deleted, which seems to be well specified empirically, the evidence against the CRRA hypothesis is very weak. Thus, the intuitive arguments of Arrow favoring CRRA may be useful.

Concluding Comments

Although much has been said about the impact of the structure of risk preferences on behavior,

virtually no empirical work examines risk structure in reduced-form demands or supplies except to test for the mere existence of risk aversion. As shown here, much more can be obtained from empirical models of risk response. Simple Euler equations that generalize the case of homogeneity can be used to refute constant relative, partial relative, or absolute risk aversion using only reduced-form equations that can be estimated econometrically using supply and demand data.

These data reject the common practice of omitting wealth data in models of producer response and are consistent with similar findings of Chavas and Holt. Policies which affect initial wealth, such as tax policies or sugar policy, will

Table 2. Test Results for CARA, CRRA, and CPRRA

Test	t and F Statistics ^a	Significance Level ^b
Model 1—with covariance term		
CARA	10.1870	<.0001
CRRA	2.1636	.0203
CPRRA	18.7675	<.0001
Model 2—without covariance term		
CARA	9.8257	<.0001
CRRA	1.2648	.1088
CPRRA	22.0638	<.0001

^a These tests are conducted at the sample means of initial wealth, price expectations, variances and covariances. Because testing for CARA is obtained by inspection of table 1, simple t -tests are used. This is asymptotically equivalent to the likelihood ratio test under our assumptions. For CRRA and CPRRA, asymptotic t and F tests, respectively, are used which are equivalent to likelihood ratio tests in large samples.

^b Significances are calculated with 24 degrees of freedom for model 1 and 25 degrees of freedom for model 2. Thus, the asymptotic tests are applied conservatively.

impact profit both directly and indirectly through initial wealth. Constant relative risk aversion has a number of modeling implications. These preferences and the log normality of wealth yields a very tractable model of choice (Hansen and Singleton). Second, risk sharing which scales wealth will have no effect on choice. Thus, denominating wealth in real or nominal terms leads to identical models of choice. Any policy change that proportionally alters final wealth similarly leaves decisions unchanged.

[Received May 1986; final revision received October 1990.]

References

- Antle, J. "Econometric Estimation of Producers' Risk Attitudes." *Amer. J. Agr. Econ.* 69(1987):509-22.
- Arrow, K. *Essays in the Theory of Risk Bearing*. Amsterdam: North-Holland Publishing Co., 1971.
- Behrman, J. R. *Supply Response in Underdeveloped Agriculture*. Amsterdam: North-Holland Publishing Co., 1968.
- Binswanger, H. "Attitudes Toward Risk: Experimental Measurement in Rural India." *Amer. J. Agr. Econ.* 62(1980):395-407.
- Briys, E., and L. Eeckhoudt. "Relative Risk Aversion in Comparative Statics." *Amer. Econ. Rev.* 75(1985):281-83.
- Chavas, Jean-Paul, and Matthew T. Holt. "Acreage Decisions Under Risk: The Case of Corn and Soybeans." *Amer. J. Agr. Econ.* 72(1990):529-38.
- Chavas, Jean-Paul, and R. Pope. "Price Uncertainty and Competitive Firm Behavior: Testable Hypotheses from Expected Utility." *J. Econ. and Bus.* 37(1985):223-35.
- Dwyer, G., Jr., and C. Lindsay. "Robert Giffen and the Irish Potato." *Amer. Econ. Rev.* 74(1984):188-92.
- Estes, E., L. Blakeslee, and R. Mittelhammer. "On the Variances of Conditional Linear Least-Squares Search Parameter Estimates." *Amer. J. Agr. Econ.* 63(1981):141-45.
- Freund, R. "The Introduction of Risk into Programming Models." *Econometrica* 24(1956):253-63.
- Friend, I., and M. Blume. "The Demand for Risky Assets." *Amer. Econ. Rev.* 65(1975):900-22.
- Hansen, L., and K. Singleton. "Stochastic Consumption, Risk Aversion, and the Temporal Behavior of Asset Returns." *J. Polit. Econ.* 91(1983):249-65.
- Hey, J. "Relative Risk Aversion in Comparative Statics: Comment." *Amer. Econ. Rev.* 75(1985):284-85.
- Hollanc, A. S. "Wage Indexation and the Effect of Inflation Uncertainty on Employment: An Empirical Analysis." *Amer. Econ. Rev.* 76(1986):235-43.
- Idaho Crop and Livestock Reporting Service. *Idaho Statistic Yearbook*. Boise ID, various years.
- Ishii, Y. "On the Theory of the Competitive Firm Under Price Uncertainty." *Amer. Econ. Rev.* 67(1977):768-69.
- Just, R. "An Investigation of the Importance of Risk in Farmers' Decisions." *Amer. J. Agr. Econ.* 56(1974):14-25.
- Just, R., and R. Pope. "On the Relationship of Input Decisions and Risk." *Risk, Uncertainty, and Agricultural Development*, ed. J. Roumasset, J. M. Boussard, and I. Singh. New York: Agricultural Development Council, 1979.
- Katz, L. "Relative Risk Aversion in Comparative Statics." *Amer. Econ. Rev.* 73(1983):452-53.
- Lau, L. "Applications of Profit Functions." *Production Economics: A Dual Approach to Theory and Applications*, ed. M. Fuss and D. McFadden. Amsterdam: North-Holland Publishing Co., 1978.
- McFadden, D. "Cost, Revenue and Profit Functions." *Production Economics: A Dual Approach to Theory and Applications*, ed. M. Fuss and D. McFadden. Amsterdam: North-Holland Publishing Co., 1978.
- Menezes, C. R., and D. L. Hansen. "On the Theory of Risk Aversion." *Int. Econ. Rev.* 11(1970):481-87.
- Pope, R. "A New Parametric Test for the Structure of Risk Preferences." *Econ. Letters* 27(1988):117-21.
- Pope, R., J.-P. Chavas, and R. Just. "Economic Welfare Evaluations for Producers Under Uncertainty." *Amer. J. Agr. Econ.* 65(1983):98-107.
- Pratt, J. "Risk Aversion in the Small and in the Large." *Econometrica* 32(1964):122-36.
- Sandino, A. "On the Theory of the Competitive Firm Under Price Uncertainty." *Amer. Econ. Rev.* 61(1971):65-73.
- Siegel, F., and J. Hoban. "Relative Risk Aversion Revisited." *Rev. Econ. and Statist.* 64(1982):481-87.
- Szpirc, G. "Measuring Risk Aversion: An Alternative Approach." *Rev. Econ. and Statist.* 68(1986):156-59.
- U.S. Department of Agriculture, Economic Research Service. *Farm Real Estate: Historical Series Data, 1950-85* (plus update). ERS Statist. Bull. No. 738, Dec. 1985.
- Wolf, C., and L. Pohlman. "The Recovery of Risk Preferences from Actual Choices." *Econometrica* 51(1983):843-50.
- Zusman, F. "An Investigation of the Dynamic Stability and Stationary States of the United States Potato Market, 1930-58." *Econometrica* 30(1962):522-47.

Factor Price Stabilization and the Competitive Firm

S. Devadoss and E. Kwan Choi

This paper investigates the effects of factor price stabilization on production decisions of the competitive firm with *ex post* production flexibility. Factor price stabilization is achieved through changes in the guaranteed minimum price and the imposed maximum price. For the risk-averse firm, the effect of factor price stabilization, through a mean-preserving contraction, on capital crucially depends on whether capital and the variable input are substitutes or complements as well as the observable characteristics of the variable input demand curve.

Key words: factor price stabilization, risk aversion.

It is well known that price uncertainty has an adverse impact on the supply of a risk-averse competitive firm which makes output decisions before observing the market price (Baron and Sandmo). Many agricultural price stabilization programs are designed to reduce price uncertainty. In the United States, the nonrecourse loan program provides downside price protection to grain producers, whereas the Farmer-Owned Reserve (FOR) program limits the upward movements of prices (Wright 1985). Similarly, the European Community price support mechanism stabilizes the grain prices between intervention and target prices (Eeckhoudt and Hansen).

The literature which studies the impacts of price uncertainty on the supply side is abundant. For example, Turnovsky showed that price uncertainty may not have a negative impact on the planned output of a risk-averse firm with *ex post* production flexibility. In the same spirit, Hartman (1976) introduced a two-factor model in which capital is a quasi-fixed input and determined *ex ante*, but the variable input is chosen after the uncertainty is resolved.

In a market equilibrium framework, price uncertainty of a primary commodity not only affects the supply side, as investigated by Baron and Sandmo, but also affects the demand side

because primary commodities are used as inputs in other industries. Subsequent literature in this area, as summarized by Robison and Barry, has examined the effects of input price uncertainty on production decisions of the competitive firm. Stewart employed a two-factor model to show that if both factors are determined *ex ante* under input price uncertainty, a risk-averse firm uses less of the risky input and more of the riskless input than a risk-neutral firm. Perrakis has shown that with *ex post* input flexibility input price uncertainty introduces a definite bias on factor demands even for risk-neutral firms.

In a recent study Wright (1984) considered a competitive firm with flexibility in both the variable input and output, and analyzed the effects of input price uncertainty on factor demands of a risk-neutral firm.¹ He showed that "the response of the quasi-fixed factor, and the expected short-run response of the variable factor, are directly related to observable characteristics of the input demand and output supply functions" (p. 443). Specifically, he showed that input price uncertainty increases (decreases) the optimal level of the quasi-fixed input of a risk-neutral firm if the slope of the variable input demand curve becomes flatter (steeper) as the quasi-fixed input increases. Wright's analysis represents an important contribution to the theory of the competitive firm with *ex post* production flexibility.

The purpose of this paper is to investigate the effects of factor price stabilization policies on

S. Devadoss and E. Kwan Choi are, respectively, an adjunct assistant professor and an associate professor, Department of Economics, Iowa State University.

Journal Paper No. J-14256 of the Iowa Agriculture and Home Economics Experiment Station Project No. 2681.

The authors are indebted to two anonymous referees for their insightful comments. The usual caveats apply.

¹ Wright also considered output price uncertainty.

input demands and output supply of the competitive firm with *ex post* production flexibility. The firm is assumed to use primary commodities (e.g., corn and wheat) as inputs. The distinguishing feature of this study is that the random input price is bounded by guaranteed minimum (floor) and imposed maximum (ceiling) prices. Factor price stabilization is implemented by a mean-preserving contraction (MPC) in the input price by changing the price floor and ceiling simultaneously. This type of MPC was first proposed by Eeckhoudt and Hansen to analyze the effect of a marginal change in output price uncertainty on supply. This approach captures the essence of stylized price stabilization policies widely adopted in the United States and the European Economic Community (EEC). For example, U.S. price support programs generally limit the movements of corn and wheat prices between the price floor (loan rate) and price ceiling (release price). Similarly, the EEC generally stabilizes domestic prices between intervention and target prices.

The paper is organized as follows. The second section presents the properties of the variable input demand and the short-run supply functions. The third section considers a risk-averse firm and examines the effects of changes in maximum and minimum prices on the optimal level of capital. The effects of an MPC on capital, expected variable input, and expected output supply are also included. The paper ends with a summary and concluding remarks.

Short-Run Behavior of the Competitive Firm

To lay the basis for analyzing the effects of input price uncertainty on factor demands, in this section we examine the short-run production decisions of the competitive firm with *ex post* input and output flexibility. The competitive firm produces a homogenous product Q , using capital input K and a variable input X . The production function $Q = F(K, X)$ is assumed monotone increasing and strictly concave in inputs in the relevant range.

Assume that capital input price, c , is known, but the price of the variable input price, r , is random with density function $g(r)$ and cumulative distribution function $G(r)$. The inputs are supplied in the domestic market. Output Q may be sold in the domestic or foreign market at the world price p . While the random input price r induces a fluctuation in the aggregate domestic

supply of Q , the domestic market is assumed to be small and the world price p is fixed and independent of the random domestic factor price r .²

The capital input is a quasi-fixed factor, i.e., K must be chosen *ex ante* before the input price r is known. In contrast, the variable input X is purchased *ex post*, i.e., after the random input price is observed. The risky input X represents a domestic factor such as corn or wheat, whose price is stabilized by commodity programs. Corn and wheat are used as inputs in the food-processing industry and also in the livestock sector. The capital input K can be viewed as plant size or number of animals placed on feed.³

The firm may engage in forward contracting, which is important in the food processing and livestock sectors. Entering the forward market does not preclude the firm from buying or selling the input in the spot market. It is well known that the separation result holds for the supplier, i.e., the output decision of the producer of X is not affected by the distribution of r or risk attitude (Danthine). However, it can be shown that the separation result will not hold for a producer who demands X as an input, because the firm has *ex post* production flexibility. Incorporating the futures contracts considerably complicates the analysis without yielding additional insights. Thus, we focus on the effects of factor price stabilization on optimal production decisions.

After the risky input price r is observed, the firm's problem *ex post* is to choose X to maximize the short-run surplus,

$$(1) \quad \pi = pF(K, X) - rX.$$

The first-order condition is $pF_X(K, X) - r = 0$. The short-run variable input demand function can be written $X = X(K, p, r)$.

We examine how the variable input demand is affected by changes in r and K . The following properties of the variable input demand function are obtained from the first-order condition:

$$(2a) \quad \partial X / \partial r = 1 / pF_{XX} < 0,$$

$$(2b) \quad \partial X / \partial K = -F_{KX} / F_{XX}.$$

Because the production function is concave in X , F_{XX} is negative and hence $\partial X / \partial r < 0$ in (2a).

² For a large economy, the world price should fluctuate with the domestic random input price. However, as Wright (p. 443) noted, the analysis of input price instability along with induced random output price considerably complicates the analysis because of the short-run production flexibility.

³ See Jarvis, and Chavas and Klemme for considering livestock animals as capital input.

K and X are complements (substitutes) if F_{KX} is positive (negative). Equation (2b) indicates that an increase in capital will shift the variable input demand curve to the right (left) if the two inputs are complements (substitutes).

Next, consider how the short-run output supply responds to changes in r and K . Substituting $X(K, p, r)$ into the production function yields the short-run supply function, $Q(K, p, r) = F[K, X(K, p, r)]$. Differentiating Q with respect to r and K gives

$$(3a) \quad \partial Q / \partial r = r / p^2 F_{XX} < 0,$$

$$(3b) \quad \partial Q(K, p, r) / \partial K = F_K - F_X(F_{KX} / F_{XX}).$$

If the two factors are complements, then $\partial Q / \partial K$ is positive. However, if the two factors are substitutes, then $\partial Q / \partial K$ can be negative. Using a cost minimization problem for a given output, Hartman (1975, p. 1290) notes that $F_K - F_X(F_{KX} / F_{XX})$ is positive (negative) if K is a normal (an inferior) factor. Henceforth, K is assumed to be a normal input.⁴

Substituting $X(K, p, r)$ into (1) gives $\pi(K, p, r) = pF(K, X(K, p, r)) - rX(K, p, r)$. This surplus function is nondecreasing in p and nonincreasing in r , convex in (p, r) and for given K it is linearly homogenous in (p, r) . The following properties of the surplus function are obtained using the envelope theorem:

$$(4a) \quad \partial \pi / \partial K = pF_K(K, X(K, p, r)),$$

$$(4b) \quad \partial \pi / \partial p = Q(K, X(K, p, r)),$$

$$(4c) \quad \partial \pi / \partial r = -X(K, p, r),$$

where $pF_K(K, X(K, p, r))$, $Q(K, X(K, p, r))$, and $X(K, p, r)$ are the value of marginal product of capital, the short-run supply, and the short-run variable input demand, respectively.

If the competitive firm employs its variable input optimally *ex post*, then the profit function is given by

$$(5) \quad \Pi(K, p, r, c) = pF(K, X(K, p, r)) - rX(K, p, r) - cK = \pi(K, p, r) - cK.$$

The results in this section will be used to analyze the impact of a marginal change in the input price uncertainty on production decisions of a risk-averse competitive firm.

Factor Price Stabilization and Risk Aversion

In this section we consider production decisions of a risk-averse competitive firm which faces the random input price r . To facilitate the exposition, we first investigate the effects of small changes in maximum and minimum prices of the variable input X on the optimal level of capital. The effects of an MPC through maximum and minimum prices on the optimal level of capital, the expected responses of the variable factor, and output supply are then examined. The results show that substitutability and complementarity between capital and the variable input play a key role in determining the effects of factor price stabilization through an MPC on the optimal level of capital of a risk-averse firm.

If there are no maximum and minimum price bounds on the random input price r , then a risk-averse competitive firm's problem is to choose K to maximize the expected utility,

$$(6) \quad \begin{aligned} \text{Max}_K \quad & EU[\Pi(K, p, r, c)] \\ &= \int_0^\infty U[\pi(K, p, r) - cK] dG(r), \end{aligned}$$

where $U(\cdot)$ is a monotone-increasing and strictly concave von Neumann-Morgenstern utility function, i.e., $U'(\Pi) > 0$, $U''(\Pi) < 0$. The first-order condition is

$$E[U' \Pi_K] = \int_0^\infty [U'(\pi_K - c)] dG(r) = 0.$$

The second-order condition requires that $EU_{KK} = E[U''(\Pi_K)^2 + U' \Pi_{KK}] < 0$.

Consider a commodity program which stabilizes the input price r . Specifically, the government intervenes in the market for X through buffer stocks so that the random input price r does not fall below a guaranteed minimum price, r_m , or rise above an imposed maximum price, r_M . This implies that r is allowed to fluctuate only between the r_m and r_M , and hence the density function $g(r)$ is truncated at r_m and r_M . In general the distribution $G(r)$ is contingent upon the aggregate supply of X . If a change in the minimum or maximum price affects the industry supply of X , it will also affect the distribution $G(r)$ (Wright and Williams). However, to focus on the firm-level analysis, we adopt the convention, as in Eeckhoudt and Hansen, that the distribution $G(r)$ is unaffected by changes in r_m or r_M .

⁴ If K represents the "size" of the firm (or number of animals), it is unlikely for a small firm to produce more output than a larger firm. We have yet to find an anomalous quasi-fixed input that is inversely related to total output.

Maximum Price

The analysis begins by considering the optimal choice of capital for a risk-averse firm when a ceiling price r_M is imposed on the random input price r . The risk-averse competitive firm's problem is to choose K to maximize the expected utility,

$$(7) \quad A^M = \int_0^{r_M} U[\pi(K, p, r) - cK]dG(r) + U[\pi(K, p, r_M) - cK](1 - G(r_M)),$$

where $(1 - G(r_M))$ is the cumulative probability that the random price r is greater than the maximum price, r_M . Using (4a), the first-order condition is

$$(8) \quad dA^M/dK = B^M(K, p, c, r_M) = \int_0^{r_M} U'[pF_K(K, X(K, p, r)) - c]dG(r) + U'_{r_M}[pF_K(K, X(K, p, r_M)) - c](1 - G(r_M)) = 0,$$

where U'_{r_M} indicates that U' is evaluated at the profit level for the sure price r_M . The function B^M indicates that dA^M/dK depends on K , p , c and r_M .

Let K^a denote the optimal level of capital of the risk-averse firm. Observe that for the first-order condition to hold, the optimum level of capital K^a satisfies the inequality,

$$(9a) \quad pF_K(K, X(K, p, r_M)) > c$$

if K and X are substitutes ($F_{KX} < 0$). Note that $F_{EX} < 0$ implies that $pF_K(K, X(K, p, r)) < pF_K(K, X(K, p, r_M))$ for $r < r_M$. The inequality in (9a) is proved by contradiction: If $pF_K(K, X(K, p, r_M)) \leq c$, then $pF_K(K, X(K, p, r)) < c$. If this were the case, (dA^M/dK) would be negative and the first-order condition (8) would not be satisfied. Thus, the inequality (9a) holds. By similar reasoning, one can verify that for the first-order condition to hold,

$$(9b) \quad pF_K(K, X(K, p, r_M)) < c$$

if K and X are complements ($F_{KX} > 0$).

Totally differentiating the function $B^M(K, p, c, r_M)$, we note that at the optimum $(\partial B^M/\partial K)dK + (\partial B^M/\partial r_M)dr_M = 0$. Solving for dK/dr_M and observing that $\partial B^M/\partial K$ is negative by the second-order condition, we get

$$\text{sign}(dK/dr_M) = \text{sign}(\partial B^M/\partial r_M).$$

To determine the sign of $(\partial B^M/\partial r_M)$, we differentiate the first-order condition (8) with respect to r_M and obtain, after rearranging terms,

$$(10) \quad (\partial B^M/\partial r_M)(1/(1 - G(r_M))) = U'_{r_M}(-X) \cdot [pF_K(K, X(K, p, r_M)) - c] + U'_{r_M} \cdot pF_{KX}(K, X(K, p, r_M))(\partial X/\partial r_M).$$

Using the results in (9a) and (9b), we see that $\partial B^M/\partial r_M$ is positive (negative) if K and X are substitutes (complements). Therefore, an increase in the maximum price r_M increases (decreases) the optimal level of capital if the two inputs are substitutes (complements).

Minimum Price

We now consider the optimal choice of capital for a risk-averse firm when a guaranteed minimum price r_m is imposed on the random input price r . The firm's problem is to choose K to maximize the expected utility

$$(11) \quad A^n = U[\pi(K, p, r_m) - cK]G(r_m) + \int_{r_m}^{\infty} U[\pi(K, p, r) - cK]dG(r).$$

The first-order condition is

$$(12) \quad dA^n/dK = B^n(K, p, c, r_m) = U'_{r_m} \cdot [pF_K(K, X(K, p, r_m)) - c]G(r_m) + \int_{r_m}^{\infty} U' \cdot [pF_K(K, X(K, p, r)) - c]dG(r) = 0,$$

where U'_{r_m} indicates that U' is evaluated at the profit level for the sure price r_m . As for the maximum price, it should be noted that for the first-order condition (12) to hold, the optimum level of capital K^a must satisfy the inequality,

$$(13) \quad pF_K(K, X(K, p, r_m)) < (>) c$$

if K and X are substitutes (complements).

To determine the effect of a change in the minimum price r_m on the optimal level of capital, we totally differentiate the function $B^n(K, p, c, r_m)$ and note that at the optimum $(\partial B^n/\partial K)dK + (\partial B^n/\partial r_m)dr_m = 0$. Because $(\partial B^n/\partial K)$ is negative, we observe that

$$\text{sign}(dK/dr_m) = \text{sign}(\partial B^n/\partial r_m).$$

To evaluate the sign of $(\partial B^n/\partial r_m)$, we differentiate the first-order condition (12) with respect to r_m and rearrange terms to obtain

$$(14) \quad (\partial B^m / \partial r_m)(1/G(r_m)) \\ = U''_{r_m} \cdot (-X) \cdot [pF_K(K, X(K, p, r_m)) - c] \\ + U'_{r_m} \cdot pF_{KX}(K, X(K, p, r_m))(\partial X / \partial r_m).$$

The sign of (dK/dr_m) is generally ambiguous because the two terms on the right side of (14) have opposite signs, regardless of whether K and X are substitutes or complements. Thus, the effect of a change in the minimum price on the optimal level of capital is generally indeterminate.

Mean-Preserving Contraction with Maximum and Minimum Prices

A marginal increase in uncertainty, examined by Sandmo, corresponds to a "stretching" of the probability distribution around a constant mean" (p. 67). Eeckhoudt and Hansen noted that "the notion of a stretching of the probability distribution has some intuitive appeal, but it is hard to think of exogenous factors which would yield this specific change in the subjective expectations of the firm" (p. 1066). They suggested that

$$(18) \quad dA/dK = B(K, p, c, r_m, r_M) = U'_{r_m} \cdot [pF_K(K, X(K, p, r_m)) - c]G(r_m) \\ + \int_{r_m}^{r_M} U' \cdot [pF_K(K, X(K, p, r)) - c]dG(r) \\ + U'_{r_M} \cdot [pF_K(K, X(K, p, r_M)) - c](1 - G(r_M)) = 0.$$

a mean-preserving spread or contraction can be implemented through changes in minimum and maximum prices.⁵ Moreover, these price bounds are widely used in commodity price stabilization programs in the United States and the European Economic Community.

We first define a mean-preserving contraction in the input price r through maximum and minimum prices around a constant mean. Let the domain of the random input price be in the open interval (a, b) such that $G(a) = 0$ and $G(b) = 1$. Consider a binding minimum price r_m ($r_m > a$) and a maximum price r_M ($r_M < b$) which preserves the mean price $E(r)$. A mean-preserving contraction or "price squeeze" satisfies the condition,⁶

$$(15) \quad \int_a^b rg(r)dr = G(r_m)r_m + \int_{r_m}^{r_M} rg(r)dr \\ + (1 - G(r_M))r_M.$$

The required change in r_M following a change in r_m to maintain the same mean can be derived by totally differentiating (15) with respect to r_m and r_M ,

$$(16) \quad dr_M/dr_m = -G(r_m)/[1 - G(r_M)] \equiv -\Theta.$$

Next, consider the optimal choice of capital for a risk-averse firm facing a set of imposed maximum and minimum prices on the random input price r . The firm's problem is to choose K to maximize the expected utility,

$$(17) \quad A = U[\pi(K, p, r_m) - cK]G(r_m) \\ + \int_{r_m}^{r_M} U[\pi(K, p, r) - cK]dG(r) \\ + U[\pi(K, p, r_M) - cK](1 - G(r_M)).$$

The first-order condition is

One can verify that, at the optimum, the following conditions hold:

$$(19a) \quad pF_K(K, X(K, p, r_m)) > c \\ > pF_K(K, X(K, p, r_M))$$

if K and X are substitutes,

$$(19b) \quad pF_K(K, X(K, p, r_m)) < c \\ < pF_K(K, X(K, p, r_M))$$

if K and X are complements.

We now analyze the effect of a mean-preserving contraction through a change in the maximum and minimum prices on the optimal level of capital K^a of the competitive firm. Consider changes in r_m and r_M , holding $E(r)$ constant. Then at the optimum,

$$(\partial B / \partial K)dK + (\partial B / \partial r_m)dr_m + (\partial B / \partial r_M)dr_M = 0.$$

Because $(\partial B / \partial K) < 0$ by the second-order condition for optimal K , we observe that using (16),

$$\text{sign}(dK/dr_m)|_{dE(r)=0} \\ = \text{sign}\{(\partial B / \partial r_m) - (\partial B / \partial r_M)\Theta\}.$$

⁵ Obviously, price stabilization through changing minimum and maximum prices is a special case of mean-preserving contraction considered by Hartman (1976) and Wright (1984).

⁶ The effect of global change in the input price uncertainty can be obtained by setting r_m and r_M equal to the mean $E(r)$.

Recall that $(-\Theta)$ is the rate of change in r_M for a small increase in r_m to maintain the same mean. We evaluate $(\partial B/\partial r_m)$ and $(\partial B/\partial r_M)$, and obtain,

$$(20) \quad (\partial B/\partial r_m) - (\partial B/\partial r_M)\Theta \\ = U''_{r_m} \cdot (-X)[pF_K(K, X(K, p, r_m)) - c]G(r_m) \\ - U''_{r_M}(-X) \cdot [pF_K(K, X(K, p, r_M)) - c]G(r_m) \\ + [U'_{r_m} \cdot pF_{KX}(K, X(K, p, r_m))(\partial X/\partial r_m) - U'_{r_M} \\ \cdot pF_{KX}(K, X(K, p, r_M))(\partial X/\partial r_M)]G(r_m).$$

Using the results in (19a) and (19b), we can unambiguously determine the signs of the first two terms on the right side of (20). Both terms are negative (positive) if K and X are substitutes (complements). The third term can be written as

$$(21) \quad pF_{KX}(K, X(K, p, r_m))(\partial X/\partial r_m) \cdot [U'_{r_m} - U'_{r_M}]G(r_m) \\ + [pF_{KX}(K, X(K, p, r_m))(\partial X/\partial r_m) - pF_{KX}(K, X(K, p, r_M))(\partial X/\partial r_M)] \cdot U'_{r_M} \cdot G(r_m).$$

Because concavity of $U(\cdot)$ implies that U'_{r_m} is less than U'_{r_M} , the first term in (21) is negative (positive) if K and X are substitutes (complements). It can be shown that the second bracketed terms in (21) are positive (negative) if π_K is concave (convex) in r .⁷ Wright (1984) noted that the curvature of π_K with respect to r indicates how a change in K affects the slope of the variable input demand curve for X . Specifically,

$$\pi_{Krr} = \pi_{rrK} = -(\partial^2 X/\partial K \partial r).$$

If π_{Krr} is positive (negative), then the slope of the variable input demand curve for X becomes flatter (steeper) as K increases. This result is formally stated below.

PROPOSITION 1. *Assume that the competitive firm is risk averse. Then, (i) an increase in the maximum price r_M increases (decreases) the optimal level of capital K^a if K and X are substitutes (complements), but the effect of an increase in the minimum price r_m is ambiguous; and*

(ii) an MPC in r decreases (increases) the optimal level of capital K^a for a risk-averse competitive firm if K and X are substitutes (complements) and if the variable input demand curve becomes flatter (steeper) as K increases.

⁷ The second bracketed terms in (21) will be positive (negative) if π_K is concave (convex) in r . Observe that $\pi_K = pF_K(K, X(K, p, r))(\partial X/\partial r)$, which is positive (negative) if K and X are substitutes (complements). Furthermore, if π_K is concave in r , i.e., $\pi_{Krr} < 0$, then $pF_{KX}(K, X(K, p, r_m))(\partial X/\partial r_m)$ is greater than $pF_{KX}(K, X(K, p, r_M))(\partial X/\partial r_M)$ regardless of whether K and X are substitutes or complements. Thus, the second bracketed terms in (21) are positive. Alternatively, if π_K is convex in r , i.e., $\pi_{Krr} > 0$, then $pF_{KX}(K, X(K, p, r_m))(\partial X/\partial r_m)$ is less than $pF_{KX}(K, X(K, p, r_M))(\partial X/\partial r_M)$ regardless of whether K and X are substitutes or complements. Thus, the second bracketed terms in (21) are negative. Also note that an expression for π_{Krr} corresponds to equation (A.7) in Wright.

It can be shown that for a risk-neutral firm, if π_{Krr} is positive (negative) then an MPC in r decreases (increases) the optimal level of capital. For a risk-averse firm, however, in addition to the property of π_{Krr} complementarity and substitutability of K and X play an important role in determining the effect of an MPC on the optimal level of capital.

Expected Variable Input Demand and Output Supply

How does an MPC in the input price r affect the expected demand for the variable factor? An MPC

affects the expected variable input demand $EX(K^n, F, r)$ directly through the change in the price bounds and also indirectly via a change in the optimal level of capital, K^a . To investigate the total effect of an MPC, let $K^{a,1}$ and $K^{a,2}$ be the optimal levels of capital of the risk-averse firm before and after an MPC, and let $X(K^{a,1}, p, r^1)$ and $X(K^{a,2}, p, r^2)$ denote the variable input demands before and after the MPC, respectively. The direct effect depends on the curvature of the variable input demand curve. Specifically if the demand curve $X(K, p, r)$ is concave (convex) in r , then for a given level of $K^{a,1}$, $EX(K^{a,1}, p, r^2) > (<) EX(K^{a,1}, p, r^1)$, i.e., an MPC in r increases (decreases) the expected demand for the variable factor.⁸

The indirect effect depends on the response of the capital input to an MPC in r , and whether

⁸ Let $H(r)$ be the induced distribution of r by truncating $G(r)$ at two end points, r_m and r_M . Thus, $h(r) = g(r)$ for $r \in (r_m, r_M)$, $h(r_m) = G(r_m)$ and $h(r_M) = 1 - G(r_M)$. The expected variable input demand can be rewritten as $E[X(K, p, r)] = \int X dH(r; \alpha)$, where $H(r; \alpha)$ is the induced distribution of r and α is a shift parameter so that an increase in α represents an MPC in the distribution H and the range of the integration is over the loose interval $(0, \infty)$. For a given value of K , differentiate EX with respect to α and integrate the result by parts twice:

$$\int_0^\infty X dH_\alpha(r; \alpha) = - \int_0^\infty X_r G_\alpha(r; \alpha) dr \\ = \int_0^\infty X_{rr} \left[\int_0^\infty G_\alpha(s; \alpha) ds \right] dr,$$

where the subscript denotes partial derivative. Note that for an MPC,

$$\int_0^\infty H_\alpha(s; \alpha) ds \leq 0.$$

Thus $X_{rr} < 0$ implies that $\int_0^\infty X dH_\alpha(r; \alpha)$ is positive.

K and X are complements or substitutes. If π_{Krr} is negative and K and X are complements ($F_{KX} > 0$), then by part (ii) of proposition 1, an MPC increases the capital input ($K^{a,2} > K^{a,1}$). In this case, the indirect effect of an MPC increases expected variable input demand, i.e., $EX(K^{a,2}, p, r^2) > EX(K^{a,1}, p, r^2)$. Alternatively, if π_{Krr} is positive and K and X are substitutes ($F_{KX} < 0$), then an MPC decreases the capital input ($K^{a,2} < K^{a,1}$). In this case, the indirect effect increases the expected variable input demand, i.e., $EX(K^{a,2}, p, r^2) > EX(K^{a,1}, p, r^2)$. Thus, the indirect effect of an MPC increases the expected variable input demand if $\pi_{Krr} \cdot F_{KX}$ is negative. However, if $\pi_{Krr} \cdot F_{KX} > 0$, then K^a may increase or decrease, and the indirect effect is ambiguous. The total effect of an MPC in r on the expected variable factor demand for the risk-averse firm is summarized below.

PROPOSITION 2. *An MPC in r increases the expected variable input demand $EX(K^a, p, r)$ if $X_{rr} < 0$, $\pi_{Krr} \cdot F_{KX} < 0$.*

Next, consider the effect of an MPC in r on the expected output supply. The direct effect of an MPC in r on expected output depends on the curvature of the supply function $Q(K^a, p, r)$ in r . Specifically, if $Q(K^a, p, r)$ is concave (convex) in r , an MPC in r increases (decreases) expected output. The indirect effect depends on how an MPC affects the optimal level of K and whether K is a normal factor. Because K is assumed to be a normal input, if $\pi_{Krr} < 0$ and $F_{KX} > 0$, then an MPC increases K^a . This in turn shifts the supply schedule $Q(K^a, p, r)$ to the right and, hence, the indirect effect increases expected output. Alternatively, if $\pi_{Krr} > 0$ and $F_{KX} < 0$, then K^a decreases and, hence, the indirect effect decreases expected output. The total effect of an MPC in r on the expected output for the risk-averse firm is summarized below.

PROPOSITION 3. *An MPC in r increases expected output $EQ(K^a, p, r)$ if $Q_{rr} < 0$, $\pi_{Krr} < 0$ and $F_{KX} > 0$ but decreases expected output if $Q_{rr} > 0$, $\pi_{Krr} > 0$ and $F_{KX} < 0$.*

Thus, the effect of an MPC on the expected output of a risk-averse firm depends also on the complementarity or substitutability between K and X .

Concluding Comments

While the literature extensively analyzed the effects of output price stabilization policies on

supply responses, the effects of factor price stabilization on production decisions have received scant attention. Price stabilization policies are applied to primary commodities such as corn and wheat which are used as inputs in the food processing and livestock sectors. Price stabilization for these commodities is achieved by the use of maximum and minimum prices in the United States and the EEC. Specifically, in the United States, price support programs generally limit the movements of corn and wheat prices between the price floor (loan rate) and price ceiling (release price) through a buffer stock scheme. Similarly, prices are stabilized between intervention and target prices in the EEC.

This paper investigates how factor price stabilization, through changes in maximum and minimum prices, affects production decisions of risk-averse firms which use primary commodities as inputs. Following Turnovsky, we emphasize the ability of the firm to adjust its output after the variable input price is known. We have shown that an increase in the maximum price increases (decreases) the optimal level of capital according to whether capital and the variable input are substitutes (complements). On the other hand, the effect of an increase in the minimum price on the optimal level of capital is generally ambiguous. The effect of an MPC in r , which requires a decrease in the maximum price and an increase in the minimum price, on the optimal level of capital of a risk-averse firm depends, in addition to the relationship between capital and the viable input, upon the observable characteristics of the variable input demand curve, i.e., how its slope changes as K changes.

[Received April 1990; final revision received October 1990.]

References

- Baron, D. "Price Uncertainty, Utility, and Industry Equilibrium in Pure Competition." *Int. Econ. Rev.* 11(1970):463-80.
- Chavas, Jean-Paul, and R. M. Klemme. "Aggregate Milk Supply Response and Investment Behavior on U.S. Dairy Farms." *Amer. J. Agr. Econ.* 68(1986):55-66.
- Danthine, Jean-Pierre. "Information, Futures Prices, and Stabilizing Speculation." *J. Econ. Theory* 17(1978):79-98.
- Eeckhoudt, L., and P. Hansen. "Minimum and Maximum Prices, Uncertainty, and the Theory of the Competitive Firm." *Amer. Econ. Rev.* 70(1980):1064-68.
- Hartman, R. "Competitive Firm and the Theory of Input

Interactions between Agricultural and Resource Policy: The Importance of Attitudes toward Risk

Howard D. Leathers and John C. Quiggin

Using a method proposed by Meyer for deriving comparative statics results in the presence of risk, this paper analyzes the effects of various agricultural and environmental policy alternatives on the choices of a risk-averse producer with a Just and Pope production function. Many commonly held beliefs about policy effects are not supported unambiguously by economic theory. For example, a tax on pesticides will not necessarily reduce pesticide use or average output, and a reduction in price of agricultural output will not necessarily lead to a reduction in use of water or agricultural chemicals.

Key words: agricultural policy, resource policy, risk.

Considerable attention has been given to the interactions between agricultural policy and environmental/resource policy (e.g., Just and Bockstael). To a great extent the policy interactions reflect the use in agricultural production of inputs, such as pesticides, fertilizer, and irrigation, that can affect environmental quality. An environmental program which reduced the use of pesticide, for example, would influence agricultural output and thereby influence the effectiveness of certain agricultural policies. Similarly, an agricultural program which increases crop production will influence fertilizer use and thereby influence the effectiveness of certain environmental policies.

This paper analyzes the effects of various agricultural and environmental policy alternatives on the choices of a risk-averse producer who uses inputs such as pesticides, chemical fertilizer, and irrigation. The analysis is based on comparative statics results for a risk-averse producer with a Just and Pope production function (1978, 1979), in which input choice can affect the mean and higher moments of output. This characteristic was designed to capture the production impacts of inputs like pesticides and fertilizer. Pesticides, for example, cushion the fall in yield that would occur during pest infestations, thus af-

fecting the variation of yield as well as average yield.

Some comparative statics results with a Just and Pope production function have been derived by Feder (1977). However, Feder does not derive all changes in output and input caused by policy-induced changes in parameters; thus, his paper is not a sufficient basis for drawing policy conclusions. What is more, despite its popularity, the usual Just and Pope production function is not an entirely satisfactory representation of an input such as pesticides.¹

The paper is organized as follows. The next section reviews the Just and Pope production function and the Meyer results for obtaining comparative statics under uncertainty. Comparative statics results then are derived for a Just and Pope production function, and the analysis is extended to the cases of truncated distributions and multiple inputs. Policy implications are considered in the concluding section.

Just and Pope Production Functions and the Meyer Results

The Just and Pope production function is $y = f(x) + h(x)\varepsilon$, where y is output, x is input, and ε is a random variable. Here f and h are assumed

Howard D. Leathers is an assistant professor, and John C. Quiggin is an associate professor, both in the Department of Agricultural and Resource Economics, University of Maryland.

¹ This point was made with unusual clarity by an anonymous reviewer, whose comments led to some of the extensions presented in the fourth section.

nonnegative. Further, ε takes the form $\varepsilon = \varepsilon^* + (\gamma) \cdot (\varepsilon^* - E(\varepsilon^*))$, where ε^* is a random variable with the same mean as ε and γ is a positive constant (Sandmo). Therefore, $E(\varepsilon^*) = E(\varepsilon) \equiv \mu_\varepsilon$, $\text{Var}(\varepsilon) = \gamma^2 \text{Var}(\varepsilon^*)$, and $\sigma_\varepsilon = \gamma \sigma_{\varepsilon^*}$, where σ_{ε^*} is the standard deviation of the random variable ε^* . This representation of ε allows a spread-preserving change in mean of ε to be indicated by a change in μ_ε , and a mean-preserving multiplicative spread of ε by a change in γ : $\partial \sigma_\varepsilon / \partial \gamma > 0$.

By assumption, increased input use increases average output at a decreasing rate: $f' + h'\mu_\varepsilon$

$$\begin{aligned} F_i(k) &\equiv \Pr(\pi_i < k) = \Pr(\varepsilon < [k + wx_i - pf(x_i)]/ph(x_i)) \\ &= \Pr(\varepsilon < \beta k / ph(x_i) + [wx_i - pf(x_i)]/ph(x_i) + \alpha / ph(x_i)) \\ &= \Pr(pf(x_i) + ph(x_i)\varepsilon - wx < \beta k + \alpha) = \Pr(\pi_j < \beta k + \alpha) \\ &= F_j(\beta k + \alpha). \end{aligned}$$

where $\beta = ph(x_i)/ph(x_i)$, and

$$\alpha = ph(x_i) \cdot [wx_i - pf(x_i)]/ph(x_i) - [wx_j - pf(x_j)]$$

> 0 , $f'' + h''\mu_\varepsilon < 0$. Some inputs (pesticides and irrigation) are risk reducing: $h' < 0$. Use of pesticide, for example, increases yield when a pest infestation occurs but does not affect yield when no infestation occurs. By eliminating the lower tail of the yield distribution, pesticide use will increase mean yield and reduce the variance of yield. Some inputs (fertilizer) are risk increasing: $h' > 0$. Use of fertilizer increases the probability of very high yields (for example, when rainfall is adequate and timely) but also increases the probability of low yields (for example, when rainfall is inadequate and chemical burning occurs). Thus, increasing fertilizer use increases both mean yield and yield variability.

A producer chooses x to maximize expected utility of income:

$$(1) \quad \max EU(\pi(x)),$$

where U is a continuously differentiable utility function defined on income or profits, π .

The comparative statics results in this paper are derived using the simplified procedure proposed by Meyer. This procedure does not require any additional assumptions about the form of the utility function or the distribution of the random variable, ε^* . The Meyer approach is valid whenever members of the choice set differ only by location and scale parameters. This section shows that the location and scale condition holds in the case of the Just and Pope production function and reviews some basic results from the Meyer paper.

Any level of the choice variable x , say x_i , cor-

responds to random variable π_i , where

$$\pi_i = pf(x_i) + ph(x_i)\varepsilon - wx_i.$$

Thus, the producer's problem is viewed as one of choosing an element of the set of random variables containing all the π_i 's as defined above.

Elements of this choice set differ only by location and scale (LS) parameters if $F_i(k) = F_j(\beta k + \alpha)$, where F_i and F_j are the cumulative density functions for random variables π_i and π_j , z is any value in the range of the *cdf*, and β and α are nonstochastic terms.

To see that this holds here:

This fact allows us to use the simplified comparative statics procedure proposed by Meyer. For the problem at hand, the solution to problem (1) above is equivalent to the solution to the problem:

$$(2) \quad \max_x V(\mu, \sigma),$$

where $\mu = pf(x) + ph(x)\mu_\varepsilon - wx$, and $\sigma = p \cdot h'(x) \cdot \sigma_\varepsilon = p \cdot h'(x) \cdot \gamma \cdot \sigma_{\varepsilon^*}$.

Meyer shows the following:

1. $V_\mu \geq 0$.
2. $V_\sigma < 0$ iff $U'' < 0$ (risk aversion).
3. $S(\mu, \sigma) \equiv -V_\sigma/V_\mu > 0$ under risk aversion.
4. If $U'' < 0$, then V is concave, so, $V_{\mu\mu} \leq 0$; $V_{\sigma\sigma} \leq 0$; and $V_{\mu\mu}V_{\sigma\sigma} - (V_{\mu\sigma})^2 \geq 0$, under risk aversion.
5. $S_\mu (<, =, >) 0$ under (decreasing, constant, increasing) absolute risk aversion (DARA, CARA, and IARA). In addition:
6. $S_\sigma \geq 0$ under DARA or CARA. (See appendix for proof.)

These results will be referred to in deriving comparative statics results in the next section.

These results can be represented using familiar graphs of indifference curves and feasible sets in (μ, σ) space, shown in figure 1.² $V(\mu, \sigma)$ is a representative indifference curve. Under risk aversion, V is upward sloping (property 3), and increases at an increasing rate (property 4).

² The graphical exposition was suggested by Kwan Choi, to whom we owe a debt of gratitude. Any errors are the responsibility of the authors.

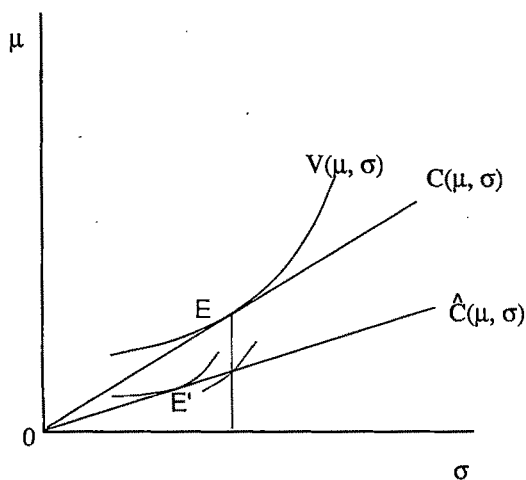


Figure 1. A graphical representation of comparative statics results

Property 5 states that, moving up along a vertical line, the indifference curves get steeper under IARA and flatter under DARA. (The DARA case is illustrated in fig. 1.) Property 6 states that moving to the right along a horizontal line, the indifference curves get steeper under DARA or CARA. The area under $C(\mu, \sigma)$ is the feasible set—the set of attainable combinations of

where

$$\mu = pf(x) + ph(x)\mu_e - wx, \text{ and}$$

$$\sigma = p \cdot h(x) \cdot \sigma_e = p \cdot h(x) \cdot \gamma \cdot \sigma_e^*.$$

One reassuringly unambiguous result from first-order conditions is

$$pf' + ph'\mu_e - w = Sph'\sigma_e^*.$$

Since $S > 0$ under risk aversion, the sign of this expression is the sign of h' , and the risk-averse producer will use more of the risk-reducing input, and less of a risk-increasing input, than will a risk-neutral producer, who sets the left-hand side equal to zero.

However, other comparative statics results are ambiguous. Comparative statics are derived from implicit differentiation of Φ . For any parameter α , $\partial x/\partial \alpha = -[\partial \Phi/\partial \alpha]/[\partial \Phi/\partial x]$. The sufficient condition for an interior optimum is that $\partial \Phi/\partial x < 0$. In the case of a risk-averse producer choosing a risk-increasing input (fertilizer), an interior solution is by no means assured. However, when the optimal choice is to use none of an input, this optimum will not be changed by small changes in parameters. Attention here is restricted to the case of interior solutions.

For interior solutions, the sign of $\partial x/\partial \alpha$ is the same as the sign of $\partial \Phi/\partial \alpha$.

$$(4a) \quad \partial \Phi/\partial w = -1 - ph'\gamma\sigma_e^* \cdot S_\mu \cdot (\partial \mu/\partial w) = -1 + ph'\gamma\sigma_e^* S_\mu x$$

$$\partial \Phi/\partial p = f' + h'\mu_e - Sh' - ph'\gamma\sigma_e^* S_\mu \cdot (\partial \mu/\partial p) - ph'\gamma\sigma_e^* S_\sigma \cdot (\partial \sigma/\partial p)$$

$$(4b) \quad = (w/p) - ph'\gamma\sigma_e^* \cdot [S_\mu \cdot (f + h\mu_e) + S_\sigma \cdot h\gamma\sigma_e^*]$$

$$(4c) \quad \partial \Phi/\partial \gamma = \left[-Sph'\sigma_e^* - ph'\gamma\sigma_e^* S_\sigma \cdot \frac{\partial \sigma}{\partial \gamma} \right] = -ph'\sigma_e^* [S + S_\sigma \gamma ph\sigma_e^*]$$

$$(4d) \quad \partial \Phi/\partial \mu_e = ph' - ph'S_\mu \cdot \frac{\partial \mu}{\partial \mu_e} = ph' [1 - S_\mu ph].$$

μ and σ . The shape of C is determined by the production function. As parameters change, the boundary of this set will change. These comparative statics are explored next.

Comparative Statics from a Just and Pope Production Function

The first-order condition associated with (2) is

$$(3) \quad V_\mu(\partial \mu/\partial x) + V_\sigma(\partial \sigma/\partial x) = 0, \text{ or} \\ (\partial \mu/\partial x) - S \cdot (\partial \sigma/\partial x) = 0; \text{ or} \\ \Phi = pf' + ph'\mu_e - w - S(\mu, \sigma)ph'\gamma\sigma_e^* = 0,$$

The comparative statics results are derived from (4) using results 1 through 6 in the above section. Comparative statics are summarized in the following propositions.

PROPOSITION 1. For a risk-averse producer and a risk-reducing input,

A. $\partial x/\partial w < 0$ under IARA, CARA (4A and result 5).

B. $\partial x/\partial p > 0$ under CARA (4B and result 5).

C. $\partial x/\partial \gamma > 0$ under DARA, CARA (4C and results 3, 6).

D. $\partial x/\partial \mu_e < 0$ under DARA, CARA (4D and result 5).

The results of proposition 1 broaden the findings of Feder with regard to pesticides. Result A confirms Feder's conclusion: "the impact of an increase in pesticide costs on pesticide dosage may be ambiguous . . . if absolute risk aversion is decreasing. . . . With constant absolute risk aversion, this ambiguity does not appear" (1979, p. 99). Here, we show that the result is also unambiguous under IARA.

The intuition behind this result can be understood by regarding each parameter change as having two effects: a direct effect and a wealth effect that operates by changing the farmer's attitude toward risk. Thus, an increase in the price of the input has the direct effect of causing the individual to buy less of the input, but the indirect effect is to reduce the individual's wealth. For a DARA individual, this reduction in wealth causes an increase in aversion to risk, which makes the individual demand more of a risk-reducing input. The balancing of these two effects creates ambiguities in the comparative statics results.

Result B shows that a similar ambiguity exists concerning the impact of an increase in output price on use of a risk-reducing input except under CARA. In order for this to be unambiguously positive, we need to have $[S_\mu(f + h\mu_e) + S_\sigma h\gamma\sigma^*] \geq 0$, which will happen if $S_\mu \geq 0$ and $S_\sigma \geq 0$. $S_\mu \geq 0$ requires IARA or CARA; $S_\sigma \geq 0$ requires DARA or CARA. The intuition here is more difficult. An increase in output price has three effects: First, it increases the expected marginal product, creating a tendency to increase input use. Second, the increase in price causes an increase in variability of revenue tending to make a producer restrict input use. As with output price risk (Sandmo, Ishii), the effect of the mean dominates the effect of variability for DARA and CARA producers. However, this is not enough to unambiguously sign the total effect. Third, the increase in price has a wealth effect, making producers more wealthy. This makes DARA producers less risk averse (encouraging less use of a risk-reducing input) and IARA producers more risk averse (encouraging more use of a risk-reducing input). Because CARA producers do not respond to a wealth effect, this third effect is irrelevant for them.

Result C states that an exogenous decrease in yield risk, holding expected yield constant, will lead to a decrease in use of risk-reducing inputs by DARA and CARA producers. The ambiguity here exists for IARA producers. This corresponds to the ambiguity for IARA producers

found in analyses of producers' response to a mean-preserving price stabilization scheme (Sandmo, Ishii).

Similarly, an exogenous decrease in μ_e , holding yield variance constant, will decrease the use of a risk-reducing input by DARA and CARA producers but not necessarily by IARA producers. This confirms Feder's result: "[Under DARA or CARA], an increase in the mean of [damage caused by a pest] will induce higher level of pesticide application" (1979, p. 99). To see that an increase in pest infestation corresponds to a decrease in μ_e , note that a drop in μ_e decreases expected yield but increases expected marginal effectiveness of input use.

PROPOSITION 2. *For a risk-averse producer and a risk-increasing input:*

A. $\partial x/\partial w < 0$ under DARA, CARA (4A and result 5)

B. $\partial x/\partial p$ is of indeterminate sign under any assumption.

C. $\partial x/\partial \gamma < 0$ under DARA, CARA (4C and results 3, 6).

D. $\partial x/\partial \mu_e > 0$ under DARA, CARA (4D and result 5).

Proposition 2 shows that ambiguities in comparative statics also exist for risk-increasing inputs. Result A shows that an increase in input price will unambiguously cause a decrease in use of risk-increasing inputs only if the producer is DARA or CARA. This is in contrast to the above result for risk-reducing inputs, where the effect was unambiguous only for IARA or CARA producers. The intuition is the same as in proposition 1. Here, the wealth effect of an input price increase makes DARA producers more risk averse and creates an additional tendency to reduce use of risk-increasing inputs.

The effect of a change in output price in the use of risk-increasing inputs cannot be determined under any assumption about attitudes toward risk. In order for this effect to be positive, it is sufficient that S_μ and S_σ be nonpositive. $S_\mu \leq 0$ under DARA, CARA, while S_σ can be nonpositive only under IARA. Note that the assumption of IARA is not sufficient to ensure that $S_\sigma \leq 0$. This fact precludes drawing unambiguous conclusions even for CARA, as in proposition 1.

An exogenous decrease in yield risk, holding mean yield constant, will increase use of a risk-increasing input by DARA and CARA producers. This differs from the result on risk-reducing inputs, where DARA and CARA producers used

less of the input in response to a mean-preserving spread of yield. An exogenous increase in mean yield, holding yield risk constant will increase use of a risk-increasing input by DARA and CARA producers. This result, too, differs from the analogous result in proposition 1, which says that DARA and CARA producers will respond to an increase in μ_e by reducing the amount of a risk-reducing input. For risk-increasing inputs, an increase in μ_e might be caused by a new seed variety, which increased expected yields, and also increased average marginal productivity of fertilizer.

These comparative statics results can be illustrated using figure 1. For example, an increase in w reduces the slope of C^3 if x is a risk increasing input ($h' > 0$). For CARA and DARA firms, indifference curves are more steeply sloped moving down a vertical line old tangency point; thus the new point of tangency must be one with lower μ and σ . This illustrates proposition 2A—an increase in w causes a decrease in use of risk-increasing inputs for DARA and CARA firms.

Model Extensions

The Just-Pope representation has a number of problems. In this section, some of the problems are identified and ways to modify the above model are suggested.

The Just and Pope production function is limited to a single input. Although in some contexts this feature could be rationalized in terms of a composite input, this approach is unattractive when the object is to analyze input demand response for a risk-reducing (or increasing) input such as irrigation or fertilizer.

The restriction may be relaxed in two ways. First, a vector z of "riskless" inputs can be added to the single risk-reducing (or increasing) input, leading to a production function of the form

$$(5a) \quad y = \phi(x, z) + h(x)\varepsilon.$$

The first-order condition on z is

$$(5b) \quad E[U'(p)(p\phi_z - w_z)] = 0,$$

where w_z is the wage rate for input z . Since the term $(p\phi_z - w_z)$ is nonstochastic, this condition can be satisfied only if $p\phi_z = w_z$. The marginal product ϕ_z depends on x but not on the stochastic term ε . Hence, it is possible to "nest"

the maximization problems, first choosing an optimal $z(x)$ satisfying $p\phi_z = w_z$, and then writing:

$$(5c) \quad f(x) = \phi(x, z(x)).$$

This approach yields the original Just-Pope representation. In order to deal with comparative static problems concerning a change in p , the slightly more complex form (5a) must be utilized. However, as long as the cross-derivative $\phi_{xz} \geq 0$, it is clear that the qualitative character of the results derived above will be unchanged.

The more general case is that of a multiple input production function in which all inputs affect risk, so that

$$(6) \quad y = f(x) + h(x)\varepsilon.$$

In the light of the analysis above, the optimization problem for (6) may be posed in the following two-stage fashion. First, for each value of σ , choose x to solve the problem

$$\max_x E\pi(x, p, w) \text{ such that } h(x)\sigma_\varepsilon = \sigma.$$

The solution to this problem determines the locus of available pairs (μ, σ) . The preference function $V(\mu, \sigma)$ may be used to determine the optimal (μ, σ) and hence the optimal x . The functional form (6) retains the linear separation, which ensures that the set of available distributions satisfies the LS condition; thus, the Meyer analysis may be applied. Input demand response is more complex, in this case, because both the preferred value of $h(x)$ and the least-cost x to yield this value must be determined. This problem will not be addressed in detail here.

The second major problem with the Just-Pope approach is that it yields a restrictive representation of risk and, particularly, of risk-reducing inputs such as pesticide. It is desirable to model these inputs as reducing losses in the worst states of the world and, hence, truncating (or, more precisely, winsorizing) the bottom tail of the distribution of outcomes. This possibility cannot be modeled in the Just-Pope framework, nor do the associated distributions satisfy the location and scale (LS) condition proposed by Meyer.

However, a number of studies (Quiggin and Anderson, Eeckhoudt and Hansen, Meyer and Ormiston 1983) have shown that risk reductions involving truncation of the bottom tail of a distribution have tractable comparative static properties. More recently, Quiggin (1986) and Meyer and Ormiston (1989) have studied a broad class of changes in risk which includes truncations and

³ The slope of C is $[\partial\mu/\partial x]/[\partial\sigma/\partial x] = (pf' + ph'\mu_e - w)/ph'\gamma\sigma_e^*$.

multiplicative spreads as special cases. Quiggin refers to such changes as monotone spreads.

The concept is best modeled by considering a random variable as a mapping g from a state space $[0, 1]$ to the outcome space \mathcal{R} such that $\omega_1 > \omega_2 \Rightarrow g(\omega_1) \geq g(\omega_2)$. Such a mapping may be regarded as the inverse of the cumulative distribution function for a continuously distributed random variable. Two variables g_1, g_2 are related by a monotone spread if $E[g_1] = E[g_2]$ and $g_1(\omega) - g_2(\omega)$ is a monotone increasing function of ω . Thus, in moving from g_1 to g_2 , the bad states get worse and the good states get better, or there is more weight in the tails of the distribution. Special cases in which two variables differ only by mean and a monotone spread include the case where g_1 is a mean-preserving multiplicative spread of g_2 (as represented by an increase in γ in the previous section), and the case where g_1 is a truncation of g_2 .

The analysis of the previous section may now be extended to cover production functions of the form

$$(7) \quad y_i = f(x_i, z_i) + \lambda(x_i)h(x^0, \omega) + (1 - \lambda(x_i))h(x^*, \omega),$$

where for two levels of input use, x^0 and x^* , the random variables $h(x^0, \omega)$ and $h(x^*, \omega)$ differ by a monotone spread, with common expected value \bar{h} .⁴

This formulation is sufficiently flexible to provide a plausible model of risk response for irrigation, fertilizer, and pesticide. To illustrate, define states of the world as the level of pest infestation, where a low state corresponds to high infestation; let x^0 represent no pesticide use, and x^* represent the level of pesticide use which eliminates pests all levels of infestation. Mapped against states, $h(x^*)$ is a horizontal line. Combined with $f(x^*)$, which is constant for all states by assumption, this gives us a constant $y(x^*)$. Mapped against states of the world, output $y(x^0)$ starts low, increases until it reaches a maximum in the state at which infestation is inconsequential, and then turns horizontal. Linear combinations of $h(x^0)$ and $h(x^*)$ all turn horizontal at the same (zero infestation) state.

Quiggin (1990b) shows that the Meyer approach can be applied to analyze comparative statics when the production function takes the form of (7). The properties of $V(\mu, \sigma)$ derived by Meyer for the LS case generalize to $V(\mu, \lambda)$.

Furthermore, since λ is linearly related to the variance and standard deviation, these properties carry over to $V(\mu, \sigma)$.

In the (7) representation, $\mu = pf(x_i, z_i) + p\bar{h} - wx - w_z \cdot z_i$, and $\sigma^2 = \lambda(x_i)\sigma_0^2 + (1 - \lambda(x_i))\sigma_*^2$, where σ_0^2 and σ_*^2 are the second central moments of the random variables $h(x^0, \omega)$ and $h(x^*, \omega)$, respectively.⁵ A change in σ_0^2 can be interpreted as a change in income variability attributable to crop insurance or any other exogenous change in variation of outcomes.

Finally, the Just-Pope and Meyer analyses have been confined to expected utility maximizers. The development of more general models of choice under uncertainty (Machina, Quiggin 1982) suggests that this is an unnecessary restriction. Quiggin (1990a) shows that the main EU comparative static results for monotone spreads carry over to the more general rank-dependent expected utility model. Quiggin (1990b) examines the relationship between preferences of the form $V(\mu, \sigma)$ and the general smooth preferences model of Machina.

Policy Implications

Although existing literature contains some of these comparative statics results, policy implications of the results have not been well developed. Many commonly held beliefs about policy effects, especially interactions between agricultural and resource policy, are not supported unambiguously by economic theory. Rather, they rest on presumptive answers to empirical questions concerning individual producer's attitudes toward risk.

In particular, the above analysis of comparative statics shows the following.

(a) A tax on pesticides will not cause all risk-averse farmers to use less pesticides, nor will it necessarily result in lower average output.

(b) A subsidy on water use will not cause all risk-averse farmers to increase water use or expected output.

Use of a risk-reducing input decreases as input price increases under the assumptions of IARA or CARA; but under the more usual assumption of DARA, an increase in input price could cause increased input use. Therefore, the standard assumptions on risk aversion are not sufficient to conclude that a water subsidy, or the failure of pesticide prices to reflect true so-

⁴ The assumption that $h(x^0, \omega)$ and $h(x^*, \omega)$ have the same expected value does not imply that expected output is the same for x^0 and x^* , since $f(x^0)$ and $f(x^*)$ can differ.

⁵ This expression for variance follows from the assumption that $h(x^0)$ and $h(x^*)$ are related by a monotone spread.

cial costs, causes agricultural pollution or creates a burden on agricultural price support programs by boosting output.

In addition, these results call into question the efficacy of Pigouvian taxes in controlling pollution derived from sources such as pesticides and irrigation. Internalizing externalities through a Pigouvian tax is straightforward when risk is not present (or when producers are risk neutral). When producers are risk averse, the tax will not only change w but will also change S in the first-order condition (3). In the case of a DARA producer, such a tax may have the perverse effect of causing input use to move farther away from the social optimum.

(c) A reduction in prices received by farmers will not necessarily lead to a reduction in use of water or agricultural chemicals.

This "expected" response holds unambiguously only for risk-reducing inputs used by CARA producers. Thus, a drop in the price received by farmers may not result in environmental improvement attributable to a decrease in use of environmentally deleterious inputs.

This result calls into question the belief that agricultural price support programs are an important cause of agricultural pollution. The theoretical results do not contradict this view, but neither do they support it. The question can only be resolved by empirical study.

(d) Introduction of a more pest resistant (or more drought resistant) seed variety, which reduces risk (σ_e) without affecting the mean of the distribution, will not necessarily result in a decline in use of agricultural chemicals.

DARA or CARA farmers will reduce use of water and pesticides but will not necessarily reduce fertilizer use. These results also hold for any other methods of reducing yield variability such as improved cultivation methods.

(e) Use of new seed varieties or production methods which decrease μ_e without affecting the variance of the distribution of yield will cause DARA and CARA farmers to use less pesticide and water but more fertilizer.

Conclusions

These results illustrate the difficulty of deriving policy effects for a risk-averse producer with a Just and Pope production function. At the extreme, theory yields no clear prediction of the effect of a change in output price on fertilizer (or other risk-increasing inputs). In order to get unambiguous and "normal" (i.e., usual, com-

monly expected) comparative statics results in every other case, it is necessary to assume constant absolute risk aversion. If one assumes decreasing absolute risk aversion, one must accept ambiguity in the impact of output and input price changes on risk-reducing inputs such as pesticides and irrigation or impose additional restrictions on risk attitudes.

In addition, exogenous changes in mean and variance of yield have opposing effects on risk-reducing and risk-increasing inputs. For example, an exogenous reduction in yield variation decreases use of pesticides but increases use of fertilizer. Thus, the overall impact of such a change on environmental damage caused by agricultural chemicals cannot be determined by theory alone.

This example illustrates the analytical difficulty highlighted by this paper: for this class of producer problems, conclusions about policy effects and prescriptions for optimal incentives cannot be based on theory alone. Rather, the analysis requires an empirical knowledge of the production function, the environmental impacts of input use, the risk attitudes of producers (the shape of each producer's utility function). Furthermore, the knowledge about the "average" attitude toward risk is not sufficient. Detailed knowledge is needed about the distribution of attitudes toward risk among farmers (Antle): how many farmers are DARA? how many are IARA? how many are "extremely" DARA? and how many are "only slightly" DARA?

In the absence of this empirical knowledge, descriptions of policy effects and interactions between agricultural and resource policy are problematic. Economic theory does not support (or contradict) judgments such as, "Price support programs cause increased agricultural pollution," or "Pollution taxes will reduce agricultural output," or "The environment would be improved by finding means other than pesticides or irrigation of reducing yield risk." This paper sounds a cautionary note to policy makers and analysts.

[Received November 1989; final revision received September 1990.]

References

- Antle, J. M. "Econometric Estimation of Producers' Risk Attitudes." *Amer. J. Agr. Econ.* 69(1987):509-23.
- Eeckhoudt, L., and P. Hansen. "Minimum and Maximum Prices, Uncertainty and the Theory of the Competitive Firm." *Amer. Econ. Rev.* 70(1980):1064-68.

- Feder, G. "Pesticides, Information, and Pest Management Under Uncertainty." *Amer. J. Agr. Econ.* 61(1979):97-103.
- . "The Impact of Uncertainty in a Class of Objective Functions." *J. Econ. Theory* 16(1977):504-12.
- Ishii, Y. "On the Theory of the Competitive Firm Under Price Uncertainty: Note." *Amer. Econ. Rev.* 67(1977):768-69.
- Just, R., and N. Bockstael, eds. *Agricultural and Resource Policy in Agricultural Systems*. New York: Springer Publishers, 1990.
- Just, R., and R. Pope. "Production Function Estimation and Related Risk Considerations." *Amer. J. Agr. Econ.* 61(1979):276-83.
- . "Stochastic Representation of Production Functions and Econometric Implications." *J. Econometrics* 7(1978):67-86.
- Machina, M. "'Expected Utility' Analysis without the Independence Axiom." *Econometrica* 50(1982):277-323.
- Meyer, J. "Two-Moment Decision Models and Expected Utility Maximization." *Amer. Econ. Rev.* 77(1987):421-30.
- Meyer, J., and M. Ormiston. "Comparative Statics of Cumulative Distribution Changes for the Class of Risk-Averse Agents." *J. Econ. Theory* 31(1983):153-69.
- . "Deterministic Transformation of Random Variables and the Comparative Statics of Risk." *J. Risk and Uncertainty* 2(1989):179-88.
- Quiggin, J. "Anticipated Utility: Some Developments in the Economic Theory of Uncertainty." Ph.D. thesis, University of New England, Armidale, Australia, 1986.
- . "A Theory of Anticipated Utility." *J. Econ. Behavior and Org.* 3(1982):323-43.
- . "Comparative Statics for Rank-Dependent Expected Utility Models." Dep. Agr. and Resour. Econ., University of Maryland, 1990a.
- . "Two Parameter Models and Expected Utility: Some Generalizations." Dep. Agr. and Resour. Econ., University of Maryland, 1990b.
- Quiggin, J., and J. Anderson, "Price Bands and Buffer Funds." *Econ. Recd.* 57(1981):67-73.
- Sandmo, A. "On the Theory of the Competitive Firm Under Uncertainty." *Amer. Econ. Rev.* 61(1971):65-73.

Appendix

Proof that $S_\sigma \geq 0$ under DARA or CARA

By result 5 above, the assumption of CARA or DARA implies $S_\mu \leq 0$. But $S_\mu = 1/V_\mu [-V_{\sigma\mu} + (V_\sigma/V_\mu)V_{\mu\mu}]$. Therefore, $[-V_{\sigma\mu} + (V_\sigma/V_\mu)V_{\mu\mu}] \leq 0$, or $F_{\sigma\mu} \geq V_\sigma/V_\mu V_{\mu\mu} \geq 0$ by results 3 and 4. Two conclusions follow from this: (A1) $V_{\sigma\mu} \geq 0$; and (A2) $V_{\sigma\mu}/V_{\mu\mu} \leq V_\sigma/V_\mu$. By result 4 (concavity of V), $V_{\sigma\sigma}V_{\mu\mu} - (V_{\sigma\mu})^2 \geq 0$, or using (A1), $V_{\sigma\sigma}/V_{\mu\mu} \leq V_{\sigma\mu}/V_\mu$. Combining this with (A2), we have $V_{\sigma\sigma}/V_{\sigma\mu} \leq V_\sigma/V_\mu$. This is sufficient to sign S_σ since $S_\sigma = 1/V_\mu [-V_{\sigma\sigma} + (V_\sigma/V_\mu)V_{\mu\sigma}]$

Joint Risk Preference-Technology Estimation with a Primal System

H. Alan Love and Steven T. Buccola

Applied studies of the firm in a risky environment have concentrated either on the firm's technology or on its risk preferences. These models result in generally inconsistent and inefficient parameter estimates. A primal model is proposed which allows a firm's preferences and technology to be estimated jointly in the presence of risk. The model is applied to Iowa corn production and estimated technology parameters are compared with those from other approaches. Modest risk aversion leads to inelastic (even backbending) per-acre supplies and input demands. Yield heteroskedasticity in inputs leads to supply heteroskedasticity in prices, especially for risk-neutral firms.

Key words: heteroskedasticity, input demands, production function, risk aversion, yields.

Econometric supply analysis has paid increasing attention to the relationship between a firm's price responses and its technology. Duality theorems, in particular, may allow one to derive mutually consistent supplies, input demands, and technology from a single cost or profit function. Unfortunately, risk and risk aversion greatly complicate the task of developing a linkage between economic choice and observed technology. Production relations must be expanded to include inputs' effects on yield variability. Dual approaches in such a context have proven troublesome because restrictions on the firm's expected utility function generally involve information about preferences, which are unknown (Pope 1982b). Attention, therefore, has turned to primal models of positive economic behavior under risk.

In the present paper, we consider alternative primal risk models which permit technology estimation only or permit input demand estimation given a technology or preference structure. Primal approaches presently in use are shown to have a number of correctable failings. Those focusing on technology alone have improperly assumed that inputs are exogenous, while those focusing on input demands have failed to estimate preference and technology parameters in a

mutually consistent manner. A model is proposed and tested which corrects these shortcomings and which provides a basis for consistent and efficient estimation of farm behavior under risk.

The paper begins with a brief review of production function and input demand models in a risky environment. Attention then is restricted to technologies in which marginal yield moments may be independent of one another. We explore implications of this technology class for input endogeneity and for joint inference of preferences and technology. Applications are provided to Iowa corn production, and our estimates are compared with those in earlier models.

Earlier Literature

Estimation of stochastic production functions received a boost when, following Harvey, Just and Pope (1979) showed that standard heteroskedasticity correction methods may be used to draw inferences about input-risk relationships. They assumed a normal yield distribution and allowed independence between an input's effects on yield risk and its effects on yield mean. The procedure involved estimating the mean portion of the production surface, fitting the risk portion through a regression on residuals, then improving efficiency of the mean estimates through Aitken-type weighting. Nelson and Preckel instead used maximum-likelihood meth-

H. Alan Love is an assistant professor, and Steven T. Buccola is a professor, Department of Agricultural and Resource Economics, Oregon State University.

The authors wish to thank David Birkes, Statistics Department, Oregon State University, for helpful comments and Heung-Dong Lee for extensive computer assistance.

ods to estimate an input's effects on the parameters of a beta yield distribution, which in turn affect the beta's mean, variance, and higher moments. Although moments of the beta distribution are not independent of one another, the model has sufficient flexibility to allow an input to increase yield mean but reduce yield variance.

Studies under risk which allow input use to respond optimally to prices generally have separated technology estimation from preference estimation. Prior estimates of technology parameters usually are imposed on input demands, leaving only risk preference terms to be estimated (Pope 1982a). A mathematical programming approach of this sort is to find optimal input allocations corresponding to alternative risk aversion levels, then to choose the risk-aversion-level-minimizing the distance between actual and optimal input rates (Wiens, Paris, Brink and McCarl).

In an alternative approach, Antle substituted prior parameter estimates of his moment-based stochastic production function into a set of input demands and used the constrained demands to fit parameters of the distribution of risk attitudes. Allowance for independence of marginal yield moments in Just and Pope's model is preserved, although the yield distribution may be nonnormal. In contrast, Loehman and Vandever have shown that an interactive biological growth model requires marginal mean and marginal variance of yield to be interdependent. They use prior estimates of the production function to simulate a parameter partially reflecting risk preference. Unfortunately, this parameter also reflects the probability moments of profit.

Separating preference from technology estimation in these ways potentially introduces both inconsistency and inefficiency. Technology parameters derived solely from production functions are inconsistent if input uses are endogenous and correlated with error terms (Pope 1982b, pp. 346-47). Whenever technology estimates are in fact inconsistent, preferences estimated conditionally on such parameters are inconsistent as well. Functional separation also fails to exploit cross-equation restrictions and error correlations that might improve efficiency.

Issues in Joint Estimation

Prospects for joint primal estimation of technology and preferences are influenced by the interdependence of marginal yield moments and

by the form of the utility function and yield distribution. For instance, interdependent marginal yield moments, as in Loehman and Vandever, imply the general production function form

$$(1) \quad Y = g(X)i(W, u),$$

where g and i are functions, Y is per-acre output, X is a vector of nonstochastic inputs, W is a vector of stochastic inputs such as rainfall, and u is a random error. A number of authors have argued for such a multiplicative model (Newbery and Stiglitz, pp. 304-5). Just and Pope's criticism of (1) is that an input's marginal effect on variance is a function of its marginal effect on mean. In addition, i must have positive skew in order to avoid possibilities of negative yields. An indeterminately large number of marginal yield moments thus enter the input demands, and joint preference-technology estimation becomes difficult.

Maintaining independence between marginal mean and variance implies

$$(2) \quad Y = f_1(X) + f_2(W) + h(X)\epsilon,$$

where $\epsilon \sim N(0, 1)$. Stochastic and nonstochastic inputs interact to affect yield's variance but not its mean. Prior to W 's occurrence, (2) has the Just-Pope form

$$(2') \quad Y = f(X) + h(X)\epsilon,$$

in which $E(Y) = f(X) = f_1(X) + E[f_2(W)]$. Despite its inconsistency with dynamic growth models and skewed yield distributions, this form is popular and holds the best potential for mutual inference of preferences and technology. We therefore concentrate on Just-Pope form (2') and, to permit tractable results, assume negative exponential utility $U(\pi) = -\exp(-\lambda\pi)$, where π is net return and λ is absolute risk aversion. Comparative statics of (2') in conjunction with nonincreasing risk aversion have been investigated by Pope and Kramer. Babcock, Chalfant, and Colander used a moment-generating function to derive specific input demands for the case of (2') and exponential utility.

Optimality Conditions

Optimal input levels for producers facing (2') are found by solving the primal problem,

$$(3) \quad \begin{aligned} \underset{x}{\text{Max}} \quad & E[U(\pi)] \\ & = \underset{x}{\text{Max}} EU[Pf(X) + Ph(X) - r'X]. \end{aligned}$$

Here P and r are (assumed known) output price and input price vector, respectively. To minimize the number of estimable parameters, we employ Cobb-Douglas forms of f and h , which in a two-variable-input case reduce (2') to

$$(4) \quad Y = AX_1^{a_1}X_2^{a_2} + BX_1^{b_1}X_2^{b_2}\epsilon,$$

with A, a_1, a_2, B, b_1, b_2 as parameters. Substituting (4) into (3), deriving first-order conditions, and using $U'(\pi)$ to denote (random) marginal utility, gives

$$(5) \quad \begin{aligned} E\{U'(\pi)[PAa_1X_1^{a_1-1}X_2^{a_2} + PBb_1X_1^{b_1-1}X_2^{b_2}\epsilon - r_1]\} &= 0, \\ E\{U'(\pi)[PAa_2X_1^{a_1}X_2^{a_2-1} + PBb_2X_1^{b_1}X_2^{b_2-1}\epsilon - r_2]\} &= 0. \end{aligned}$$

Taking expectations and dividing by $E[U'(\pi)]$ equates certainty equivalents of value marginal products with their corresponding input prices:

$$(6) \quad \begin{aligned} PAa_1X_1^{a_1-1}X_2^{a_2} + PBb_1X_1^{b_1-1}X_2^{b_2}t &= r_1, \\ PAa_2X_1^{a_1}X_2^{a_2-1} + PBb_2X_1^{b_1}X_2^{b_2-1}t &= r_2, \end{aligned}$$

where $t = E[U'(\pi)\epsilon]/E[U'(\pi)]$.

Parameter t is a function of both risk and risk aversion. If profit is expressed as $\pi = \mu + \sigma\epsilon$, in which μ is profit's mean and σ is its standard deviation, then t 's denominator is $E[U'(\pi)] = \lambda E[\exp(-\lambda\pi)] = \lambda E[\exp(-\lambda\mu - \lambda\sigma\epsilon)]$, that is λ times the expectation of a lognormally distributed variate with parameters $-\lambda\mu, \lambda^2\sigma^2$. The latter expectation is just $E[\exp(-\lambda\mu - \lambda\sigma\epsilon)] = \exp(-\lambda\mu + \lambda^2\sigma^2/2)$ (Johnson and Kotz). In like manner, t 's numerator becomes $E[U'(\pi)\epsilon] = \lambda E[\exp(-\lambda\mu - \lambda\sigma\epsilon)\epsilon] = [\lambda \exp(-\lambda\mu)] E[\exp(-\lambda\sigma\epsilon)\epsilon] = [\lambda \exp(-\lambda\mu)][-\lambda\sigma \exp(\lambda^2\sigma^2/2)] = -\lambda^2\sigma \exp(-\lambda\mu + \lambda^2\sigma^2/2)$.¹ Dividing numerator by denominator gives $t = E[U'(\pi)\epsilon]/E[U'(\pi)] = -\lambda\sigma$. The appearance of absolute risk aversion parameter λ in (6) only as a multiple of profit risk σ implies the two terms never can be separately identified in the input demands without introducing information from production function (4).

From (2'), (3), and (4), profit standard deviation $\sigma = P[\text{Var}(Y)]^{1/2} = PBX_1^{b_1}X_2^{b_2}$. Substituting the latter and $t = -\lambda\sigma$ into (6) and combining terms results in

$$(6'a) \quad PAa_1X_1^{a_1-1}X_2^{a_2} - P^2\lambda B^2b_1X_1^{2b_1-1}X_2^{2b_2} = r_1,$$

$$(6'b) \quad PAa_2X_1^{a_1}X_2^{a_2-1} - P^2\lambda B^2b_2X_1^{2b_1}X_2^{2b_2-1} = r_2.$$

Equations (6') are a specialization of Pope and Kramer's (p. 492) first-order conditions and differ from the Babcock-Chalfant-Collender input demands only in suppressing acreage terms. Inasmuch as $\lambda\sigma^2/2 = P^2\lambda B^2b_1X_1^{2b_1-1}X_2^{2b_2}/2$ is the producer's risk premium (RP), the marginal risk premium (MRP) with respect to the first input (MRP_1) is $\delta RP/\delta X_1 = P^2\lambda B^2b_1X_1^{2b_1-1}X_2^{2b_2}$, precisely the second right-hand side term in (6'a). MRP_2 is similarly given in (6'b). Pope and Kramer, and MacMinn and Holtmann, show that a

risk averter's optimal use of both inputs exceeds (falls short of) that in the riskless or risk-neutral case if the inputs are complementary and if $\delta h/\delta X_1, \delta h/\delta X_2 < 0 (> 0)$. In (4), these conditions are satisfied if $b_1, b_2 < 0 (> 0)$, that is, if yield is heteroskedastic such that yield variance decreases (increases) as inputs increase. Clearly, input demands need not shift parallel when λ changes; in fact, the demands may become positively sloped if risk aversion is high enough (Pope and Kramer, pp. 495-96).

Input Demand-Production System

All the parameters in (6') are in principle separately identifiable. Feasibility of estimating any parameter, including product λB^2 , on the basis of (6') alone depends on the associated demand error structure. Unfortunately, production function (4) provides no information about this structure. One approach is to solve for the input demands, then add an error to each in order to represent farmers' random deviations in identifying the true optimum. However, closed-form expressions for the optimal inputs do not exist. Solving (6') as far as possible for the optimal variable factors gives

$$(7) \quad X_1 = [r_1/(PAa_1X_2^{a_2} - P^2\lambda B^2b_1X_1^{2b_1-a_1}X_2^{2b_2})]^{1/(a_1-1)},$$

$$X_2 = [r_2/(PAa_2X_1^{a_1} - P^2\lambda B^2b_2X_1^{2b_1}X_2^{2b_2-a_2})]^{1/(a_2-1)},$$

so that optimal factors can only be determined simultaneously. Full information about a factor's optimum cannot be determined from price levels.

Simultaneity in (7) ensures that the error in each equation is generally correlated with all input choices even if optimization errors are spec-

¹ $E[\exp(-\lambda\sigma\epsilon)\epsilon] = -\lambda\sigma \exp(\lambda^2\sigma^2/2)$ can be proven by noting that the moment generating function of ϵ is $m(-\lambda\sigma) = E[\exp(-\lambda\sigma\epsilon)] = \exp(\lambda^2\sigma^2/2)$. Differentiating with respect to $-\lambda\sigma$ gives $m'(-\lambda\sigma) = E[\exp(-\lambda\sigma\epsilon)\epsilon] = -\lambda\sigma \exp(\lambda^2\sigma^2/2)$.

ified additively. As an alternative to additive errors, suppose that optimization mistakes occur in the form of random failure to satisfy first-order conditions (6'). Assuming choices are optimal on average, errors v_1, v_2 are added to (6'a) and (6'b) such that $E(v_1) = E(v_2) = 0$ and $\text{Var}(v_1) = S_1^2, \text{Var}(v_2) = S_2^2$.

Estimating production function (4) jointly with demand system (6') has several advantages. First, the two equation sets contain common parameters, so joint estimation with cross-equation restrictions can improve estimation efficiency compared to the use of (6') alone. Second, considering contemporaneous error correlations between the equation sets also can improve efficiency. Third, including (4) is necessary—and sufficient—for separately identifying absolute risk aversion λ and risk coefficient B^2 in (6').

To make clear how (4) may be used for these purposes, observe that the heteroskedasticity in its error can be removed by dividing each term by $X_1^{b_1} X_2^{b_2}$. The result is

$$(4') \quad YX_1^{-b_1} X_2^{-b_2} - AX_1^{a_1-b_1} X_2^{a_2-b_2} = B\epsilon,$$

with homoskedastic error $B\epsilon \sim N(0, B^2)$. Equations (4') and (6') now may be fitted jointly once a covariance structure for residuals $v_1, v_2, B\epsilon$ has been specified. Substituting into (4') consistent estimates of A, a_i, b_i that were derived from simultaneously fitting (4') and (6') gives $B\epsilon^*$ a consistent estimate of the production error vector. Log absolute value of $B\epsilon^*$ is $\ln|B\epsilon^*| = \ln B + \ln|\epsilon^*|$, the expectation of which is $\ln B + E \ln|\epsilon^*| = \ln B - 0.6352$ (Buccola and McCarl, p. 734). Adding 0.6352 and exponentiating gives a consistent estimate of B , which when squared and divided into the consistent estimate of λB^2 from (4') and (6') gives a consistent estimate of λ .

Inasmuch as inputs X in (4') are endogenous, fitting (4') along with (6') equivalently estimates the risk-averse firm's per-acre supply (yield) function

$$(8) \quad Y^* = f(X^*) + h(X^*)\epsilon,$$

where X^* is the vector of optimal input levels. Since, through (6'), X^* depends on prices P and r , supply must be heteroskedastic in the prices whenever yield is heteroskedastic in inputs.

Estimation Procedures

We here fit an extension of system (4'), (6') to corn-soybean farm data in Iowa. Resulting pro-

duction function estimates are compared with (a) those derived with the method of Just and Pope and (b) those estimated by Nelson and Preckel with the conditional beta distribution. Hausman's specification test is used to check for independence between input use and stochastic production error in the Just-Pope model. We jointly estimate the aggregate risk aversion parameters, contrast risk-averse with risk-neutral supplies and input demands, and discuss implications for supply heteroskedasticity.

The yield and input data are taken from Nelson and Preckel. They consist of farm-level information on per-acre corn yield; nitrogen, phosphorus, and potassium application; soil slope and clay content; and two dummy variables indicating, respectively, whether a legume preceded the corn crop by one or two years. Data were drawn from Linn ($n = 103$), Muscatine ($n = 55$) and Fayette ($n = 106$) Counties, Iowa, from 1954 through 1969.² Observations on a given farm are included only for years in which corn was grown. Because some farms grew corn sporadically during the study period, cross-farm differences in parameters or error variances cannot be adequately considered.

Substantial changes in both real prices and farm input proportions occurred during 1964–69, so it is a suitable period for testing hypotheses about price-input relations. Iowa input prices were taken from annual issues of U.S. Department of Agriculture's *Agricultural Prices and Agricultural Resources: Situation and Outlook Report*. Corn price was the average of high and low March closing prices of the Chicago corn futures contract maturing the following September. All prices were deflated with the Consumer Price Index (CPI).

Of the inputs for which data were available, only nitrogen, phosphorus, and potassium are purchasable in the medium run. Hence, equations (6) were extended to a system of three input demands. Soil slope and clay and the two dummy variables were included as exogenous factors in each equation, along with output and input market prices P, r . This extended version of (4'), (6') was estimated with nonlinear 3SLS combining the indicated cross-equation restrictions on parameters and holding λB^2 fixed at trial

² The data sets, collected by the Iowa Agricultural Experiment Station, are described in more detail in Nelson and Preckel. We wish to thank Carl Nelson for kindly furnishing them to us. Original data included soil type, insecticide use, and planting date. However, because many observations on these variables were missing, they could not be employed effectively in either the Nelson-Preckel study or our own.

values (Hall, Schnake, and Cummins). Instrumental variables employed in the first nonlinear three-stage least squares (NL3SLS) estimation stage at each trial value included all the exogenous variables plus soybean futures prices, which should be correlated with corn profitability and input usage but unrelated to individual farms' error terms.

A grid search, finally, was conducted on alternative λB^2 levels and the parameter set selected which minimized system sum square errors. Substituting optimal A , a_i , and b_i estimates back into (4') gave consistent estimates of $|B\epsilon|$, the adjusted mean log of which provided a consistent estimate of B and, through the NL3SLS estimate of λB^2 , a consistent estimate of λ . The NL3SLS estimator provides, in addition, consistent estimates of all the other parameters regardless of the underlying distribution of random errors v_1, v_2 (Amemiya, Gallant, Jorgenson and Laffont). Our grid search on λB^2 implies that the usually estimated standard errors are consistent conditional on the optimized λB^2 value.

Just-Pope-type estimates of (4) were derived by applying nonlinear least squares (NLS) to obtain stage 1 estimates of A and the a_i 's. Logs of absolute values of the residuals $u = BX_1^{b_1}X_2^{b_2}$ then were regressed against $\ln|BX_1^{b_1}X_2^{b_2} \dots \epsilon|$ to derive stage 1 estimates of B and the b_i 's. A stage 2 estimate of A and the a_i 's was obtained by applying NLS to the weighted regression (4') and a stage 2 estimate of B and the b_i 's was derived by repeating the log linear regression on residuals.³ A Just-Pope-type routine provides not only a point of comparison for the NL3SLS system estimates but a useful set of starting values for these estimates as well.

If, as (6') and (7) imply, inputs X are endogenous, they should be correlated with error term ϵ . The Just-Pope parameter estimates are then inconsistent (Just and Pope 1978, pp. 75–77). On the other hand, stochastic dependence may be so small that insignificant bias results. To implement the Hausman test of the null hypothesis, $H_0: X$ independent of ϵ in (4'), we construct two sets of estimates of the technology parameters: (a) the Just-Pope ($\hat{\beta}_o$), which is consistent, asymptotically normal, and asymptotically efficient when H_0 is true but inconsistent when H_0 is false; and (b) a nonlinear two-stage least squares (2SLS) version of the Just-Pope ($\hat{\beta}_a$),

which is asymptotically inefficient if H_0 is true and consistent and asymptotically normal whether it is true or false (Sargan, pp. 54–60; White).⁴ The appropriate test statistic is asymptotically chi-square with degrees of freedom equal to the number of elements in $(\hat{\beta}_a - \hat{\beta}_o)$ (Hausman, p. 1263). It was computed for each of the three Iowa counties, focusing only on yield-mean effects $f(X)$ because consistent estimates of the corresponding asymptotic covariance matrices $\hat{V}(\beta_o)$, $\hat{V}(\beta_a)$ are readily available.

Results

Results of the system estimates for Linn County are shown in the right column of table 1, alongside the Nelson-Preckel (pp. 375–76) and Just-Pope-type estimates in the left and center columns. (Muscatine and Fayette County results are available on request.) Dummies signifying previous legume crops were not significant in any of the regressions and were dropped. A wide range of values for λB^2 was tried for each county; in all cases, the optimal values fell in the positive range. NL3SLS estimates of technology parameters A , a_i , and b_i were only modestly sensitive to parameterized λB^2 levels in the neighborhood of their optima. Primal system t -values given in parenthesis in table 1 are conditional on the λB^2 chosen. Bracketed t -values are derived from estimates of the corresponding unconditional standard errors.⁵

Parameter Estimates

System estimates of the inputs' yield mean effects were fairly close to the Just-Pope (JP) estimates and in absolute value somewhat below the Nelson-Preckel (NP) figures. For example, whereas NP estimate that a 10% increase in potassium application in Linn County increased mean corn yield by 13%, the primal system and Just-Pope models show mean yield rising by only 3% and 5%, respectively. Phosphorus' curious negative effect on mean yield in NP's Linn

³ With small samples, this last log linear regression likely improves bias in B estimates and efficiency in a_i estimates (Buccola and McCarl, pp. 735–37). $\ln(B)$ also was adjusted by +0.6352 to account for residual asymptotic bias resulting from the fact that $E|\epsilon| = -0.6352$ in each log linear regression.

⁴ Estimator $\hat{\beta}_a$ is developed identically to $\hat{\beta}_o$ except that inputs X specified at each Just-Pope stage are predicted rather than actual values. The IV model utilized for this employed the same instruments as were used in the NL3SLS model of (4'), (6').

⁵ Unconditional standard errors were estimated by numerically calculating the system's gradient vector $\partial f(\beta)/\partial \beta$, then constructing the covariance matrix given in Amemiya (p. 963). A standard error of λB^2 was obtained in this manner, but separate standard errors for λ and B^2 could be derived only if the covariance and distributions of these separate terms were known.

Table 1. Elasticity of Corn Yield Mean and Corn Yield Standard Deviation with Respect to Selected Inputs, Linn County, Iowa, 1964–1969

Input	Conditional Beta (Nelson and Preckel) ^a	Just-Pope Method	Primal System
<u>Yield Mean</u>			
A ^c		48.85 (3.89) ^b	40.36 (4.19) [3.86]
Nitrogen (N)	0.46	0.05 (2.40)	0.02 (1.41) [6.01]
Phosphorus (P)	−0.35	−0.01 (−1.61)	0.07 (2.05) [4.36]
Potassium (K)	0.13	0.05 (0.91)	0.03 (1.90) [4.33]
Slope	−0.06	0.04 (1.14)	0.07 (1.97) [1.66]
Clay	0.99	0.18 (2.87)	0.21 (4.20) [2.73]
<u>Yield Standard Deviation</u>			
B ^c		4.87	0.21
Nitrogen (N)	0.29	0.06 (0.59)	0.09 (1.64) [3.27]
Phosphorus (P)	0.19	0.56 (1.71)	0.80 (6.67) [8.11]
Potassium (K)	−0.22	−0.22 (−0.84)	0.35 (4.12) [7.03]

^a Values for the mean were derived from Nelson and Preckel's table 3. Values for the standard deviation are one-half of Nelson and Preckel's table 5 figures.

^b Numbers in parentheses are *t*-values. In the primal system, parenthesized *t*-values are conditional on the optimal λB^2 value; bracketed *t*-values are unconditional (see text). Sample size is 103.

^c Yields are measured in bushels per acre, fertilizers in pounds per acre. Sample means were N = 83.1, P = 50.1, K = 48.8, slope = 3.3, clay = 22.1, yield = 122.7.

County model contrasts with a significantly positive effect in the system model.

All three models implied, in each county investigated, strong fertilizer impacts on yield risk. For example, raising the nitrogen application rate by 10% in Linn County increased yield standard deviation by about 1% (thus variance by 2%) in the system model and by about 3% in Nelson and Preckel's conditional beta model. Both the NP and JP models suggested some fertilizers are risk reducing (in Linn, the risk reducer is potassium). All fertilizers in the system model are, in contrast, risk increasing or nonsignificantly affect risk. Risk-portion estimates tended to differ widely among the three approaches, suggesting risk effects are more difficult to estimate accurately than are mean effects.

Substituting the primal system elasticities into (4'), taking logs of the absolute values of the resulting time series, adding 0.6352 to the expectations of the series, then exponentiating gave *B* estimates of 0.209, 1.353, and 2.031 for Linn, Muscatine, and Fayette, respectively. Combining these consistent estimates of *B* with the system estimates of λB^2 generated an absolute risk aversion (*A*) estimate of 0.016 for Linn, 0.538

for Muscatine, and 0.140 for Fayette County (1967 dollars per-acre basis). These are consistent with the common assumption that producers tend to be risk averse.

Quasi-likelihood ratio tests (Gallant and Jorgenson) were performed using the unconditional covariance estimates to determine the significance of differences between Just-Pope estimates and those obtained using the NL3SLS primal system. The null hypothesis that the mean-yield elasticities obtained from the two estimators are the same was rejected at the 0.05 level for all counties. Calculated test statistics were 869.60, 26.15, and 31.65 for Linn, Muscatine, and Fayette compared to a tabled chi-square value of 11.07 at the 5% test level with five degrees of freedom. The null hypothesis that variance-of-yield elasticities were the same between the two estimators also was rejected at the 0.05 level in every case. Statistics were 1365.23, 286.53, and 9547.71 compared to a tabled chi-square value of 7.31 at the 5% test level with three degrees of freedom. Recall, however, that differences in the two models' yield mean effects are much smaller than are differences in their yield risk effects.

Calculated values of Hausman's specification error statistic were 6.52, 2.39, and 3.05 for the three counties, compared to a critical value of 11.07 at the 5% level with five degrees of freedom. The null hypothesis of independence between input usage and the production function residual cannot be rejected at a reasonable confidence level. The Just-Pope approach thus provides reasonably consistent estimates of the production function's mean-portion parameters in the cases examined. Yet, these estimates still are less efficient than the system estimates because they fail to take account of cross-equation parameter restrictions and error correlations.

Per-Acre Supply and Input Demands

Risk aversion and inputs' effects on yield risk combine to have significant implications for per-acre supplies and input demands. Implications for Linn County are illustrated by substituting the system estimates of production function and risk aversion parameters into (6') and calculat-

ing per-acre fertilizer demands. The latter are then substituted into (8) to determine means and standard deviations of per-acre corn quantities supplied.⁶

The risk-neutral grower's potassium demand curve, which ignores potassium's positive effect on yield risk, ranged from 90 pounds per acre at 5.0¢ per pound to 140 pounds per acre at 3.5¢ per pound (1967 dollars). When risk aversion ($\lambda = 0.016$) was taken into account, marginal risk premium turned positive; demand shifted leftward and passed nearly vertically through the sample mean use rate of 48.8 pounds of potassium per acre. Linn's corn price-potassium demand relations are shown in figure 1. Here, introducing risk aversion not only steepens demand but causes it to backbend, as higher output prices exacerbate the marginal risk effect of any input increases. Possible backbending of the output price-input demand relation is one reason com-

⁶ MathCAD 2.0 was used to find numerical solutions of the optimal inputs because (6') cannot be expressed in explicit form.

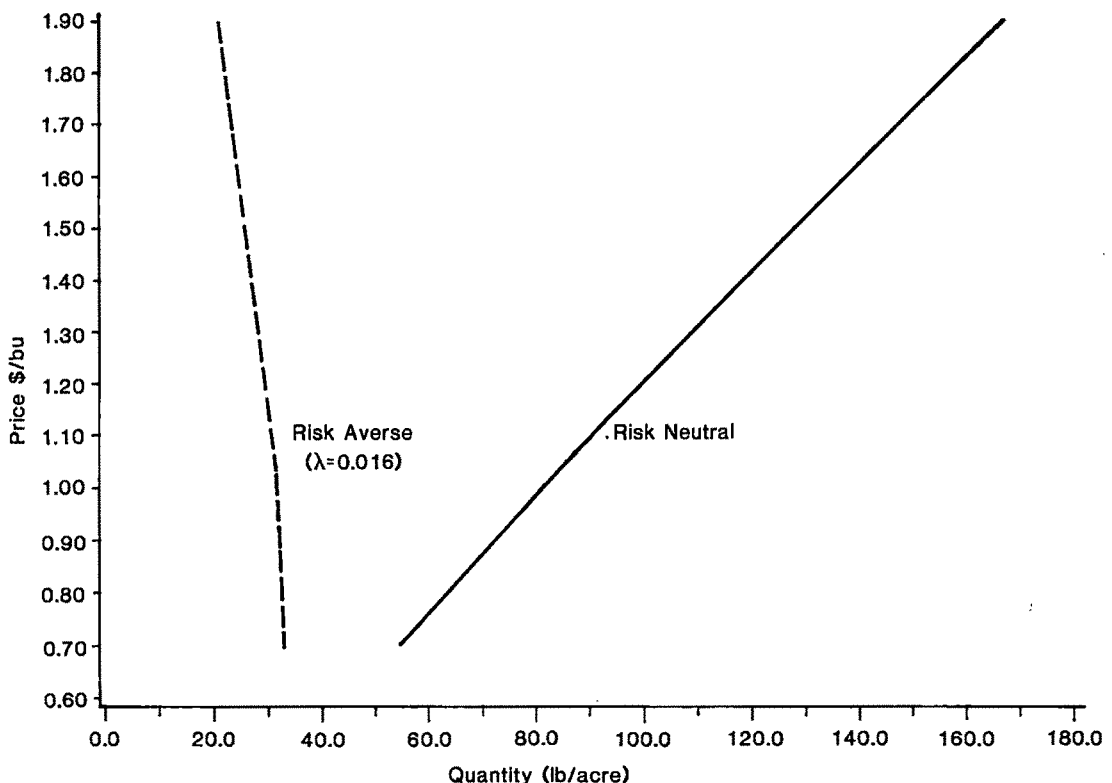


Figure 1. Per-acre potassium demand as a function of corn prices, Linn County, Iowa, 1964-69

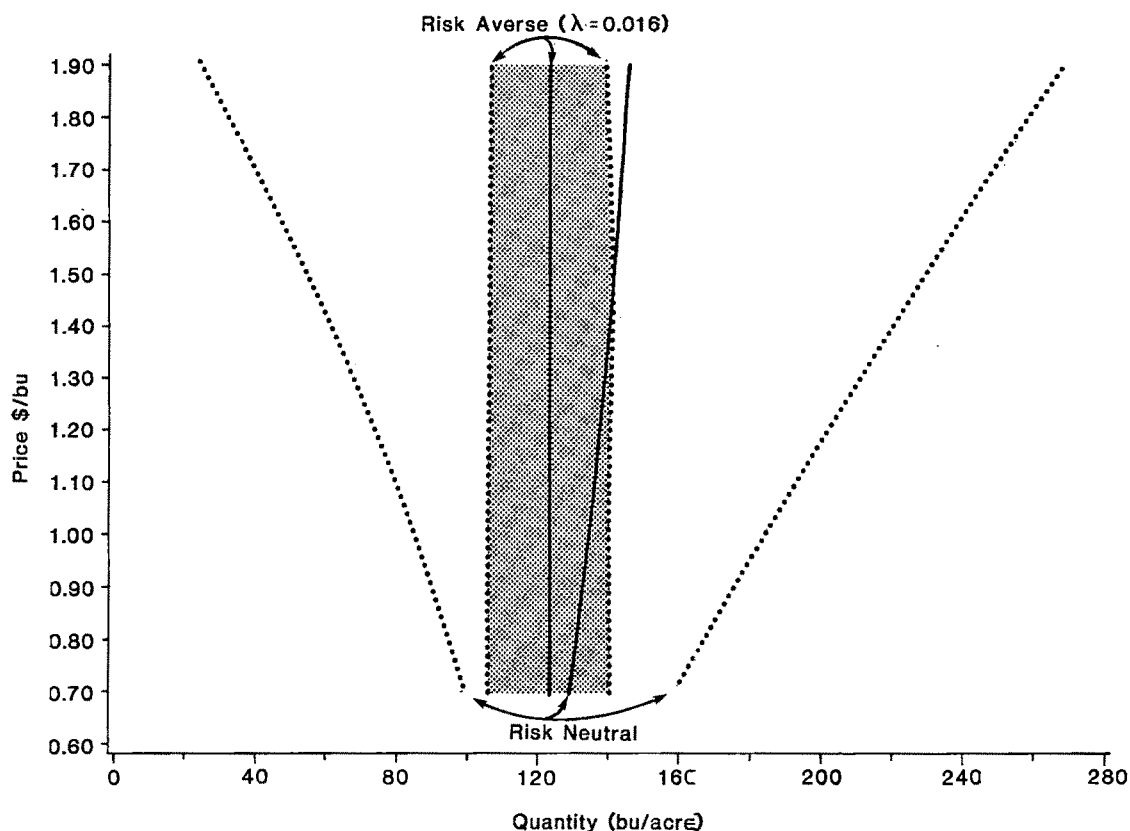


Figure 2. Mean and standard deviation of per-acre corn yield as a function of corn prices, Linn County, Iowa, 1964–69

parative statics of the risk-averse firm are so difficult to generalize (Pope and Kramer).

Finally, figure 2 gives per-acre corn supplies at various levels of expected corn price. Solid lines in the figure indicate supply expectations, while dotted lines indicate one-standard-deviation confidence intervals around these expectations. (The confidence interval for the risk-averse farmer is shaded, and the interval for the risk-neutral farmer is left unshaded.) A risk-neutral grower's relatively elastic fertilizer demands result in a slight increase in per-acre expected supply as corn price rises (right-hand solid line in fig. 2). In contrast, the risk averter's inelastic fertilizer demands combine with low mean-yield coefficients (top, table 1) to produce relatively inelastic corn supply relations (left-hand solid line in fig. 2). In fact, per-acre supply backbends slightly, reflecting the high and increasing rate at which increases in the known output price augment profit risk. Just and Zilberman also have shown that negative supply elasticities can occur under constant absolute risk aversion,

even when partial risk aversion is as low as 0.60. Average partial risk aversion for the sample of Linn County farmers was 2.34 at sample means.

The dotted confidence intervals in figure 2 indicate the extent to which yield's heteroskedasticity in inputs (bottom, table 1) leads to supply heteroskedasticity in price. Risk averters' steep input demands mean that corn price changes will little affect yield risk. For these farmers, yield variance is slightly lower at high corn prices than at low corn prices because some input uses are slightly lower at high prices as well. Risk neutrality, on the other hand, encourages greater input demand elasticity and hence a per-acre supply variability that reacts more strongly to output prices.

Conclusions

Farm technology parameters usually are assessed independently of the risk preferences of farm operators. Such separation raises practical

difficulties about sample comparability and fails to take maximum advantage of available information. We have shown that a primal system consisting of the input demands and production function may indeed be specified through which mean production, risk, and risk preference parameters can be estimated jointly. The system approach allows assessment of farmers' revealed opinions about production function relationships, including inputs' marginal impacts on yield risk. It can be used, further, to derive utility or risk preference estimates that are mutually consistent with these revealed opinions. If the model is correctly specified, consistency and efficiency generally improve compared to separate technology and utility analysis.

The approach is not without cost. Extending such a system to nonnormally distributed errors or nonconstant risk aversion is not straightforward. Incorporating nonnormal distributions into a system context probably would require using a particular nonnormal family, for example the conditional beta. Consideration of price risk also would complicate the model, and the present framework may give inaccurate risk aversion estimates in cases where price risk is substantial.

Primal system estimates of yield mean elasticities were fairly close to (but significantly differed from) those derived with a Just-Pope-type estimator. Elasticities of yield risk often differed substantially from the Just-Pope-type estimates, falling in several cases closer to the Nelson-Preckel figures. Apparently modest yield risk elasticities led to rather inelastic per-acre supplies and input demands even when risk aversion was moderate. This does not mean total supply and input demand would be inelastic as well, inasmuch as acreage may remain significantly elastic. It does, however, help explain the comparative rigidity of farm input proportions, which often is ascribed to output-maximizing behavior or asset fixity.

[Received February 1990; final revision received October 1990.]

References

- Amemiya, T. "The Maximum Likelihood and the Nonlinear Three-Stage Least Squares Estimator in the General Nonlinear Simultaneous Equation Model." *Econometrica* 45(1977):955-68.
- Antle, J. M. "Econometric Estimation of Producers' Risk Attitudes." *Amer. J. Agr. Econ.* 69(1987):509-22.
- Babcock, B. A., J. A. Chalfant, and R. N. Collender. "Simultaneous Input Demands and Land Allocation in Agricultural Production Under Uncertainty." *West. J. Agr. Econ.* 12(1987):207-15.
- Brink, L., and B. A. McCarl. "The Tradeoff between Expected Return and Risk among Cornbelt Farmers." *Amer. J. Agr. Econ.* 60(1978):259-63.
- Buccola, S. T., and B. A. McCarl. "Small-Sample Evaluation of Mean-Variance Production Function Estimators." *Amer. J. Agr. Econ.* 68(1986):732-38.
- Gallant, A. R. "Three-Stage Least Squares Estimates for a System of Simultaneous, Nonlinear, Implicit Equations." *J. Econometrics* 5(1977):71-88.
- Gallant, A. R., and D. W. Jorgenson. "Statistical Inference for a System of Simultaneous, Nonlinear, Implicit Equations in the Context of Instrumental Variable Estimation." *J. Econometrics* 11(1979):275-302.
- Hall, B., R. Schnake, and C. Cummins. *Time-Series Processor Version 4.1*. Palo Alto CA: TSP International, 1987.
- Harvey, A. C. "Estimating Regression Models with Multiplicative Heteroskedasticity." *Econometrica* 44(1976):461-65.
- Hausman, J. A. "Specification Tests in Econometrics." *Econometrica* 46(1978):1251-71.
- Johnson, N. L., and S. Kotz. *Continuous Univariate Distributions*. Boston: Houghton-Mifflin Co., 1970.
- Jorgenson, D. W., and J. Laffont. "Efficient Estimation of Nonlinear Simultaneous Equations with Additive Disturbances." *Ann. Econ. and Soc. Measure.* 3(1974):615-40.
- Just, R. E., and R. D. Pope. "Production Function Estimation and Related Risk Considerations." *Amer. J. Agr. Econ.* 61(1979):276-84.
- . "Stochastic Specification of Production Functions and Economic Implications." *J. Econometrics* 7(1978):67-86.
- Just, R. E., and D. Zilberman. "Does the Law of Supply Hold Under Uncertainty?" *Econ. J.* 96(1986):514-24.
- Loehman, E., and M. Vandever. "Testing Efficiency of Production Choices Under Risk Using the Stochastic Production Function." Paper presented at the AAEA annual meeting, Knoxville TN, 31 July-3 Aug. 1988.
- MacMin, R. D., and A. G. Holtmann. "Technological Uncertainty and the Theory of the Firm." *S. J. Econ.* 50(1983):120-36.
- Nelson, C. H., and P. Preckel. "The Conditional Beta Distribution as a Stochastic Production Function." *Amer. J. Agr. Econ.* 71(1989):370-78.
- Newbery, D. M. G., and J. E. Stiglitz. *The Theory of Commodity Price Stabilization*. Oxford: Clarendon Press, 1981.
- Paris, Q. "Revenue and Cost Uncertainty, Generalized Mean-Variance, and the Linear Complementarity Problem." *Amer. J. Agr. Econ.* 61(1979):268-75.
- Pope, R. D. "Empirical Estimation and Use of Risk Preferences: An Appraisal of Estimation Methods That Use Actual Economic Decisions." *Amer. J. Agr. Econ.* 64(1982a):376-83.
- . "To Dual or Not to Dual." *West. J. Agr. Econ.* 7(1982b):337-51.

- Pope, R. D., and R. A. Kramer. "Production Uncertainty and Factor Demands for the Competitive Firm." *S. J. Econ.* 46(1979):489-501.
- Sargan, D. *Lectures on Advanced Econometric Theory*. Oxford: Basil Blackwell, 1988.
- U.S. Department of Agriculture, Economic Research Service. *Agricultural Prices* (Annual Summaries). Washington DC.
- . *Agricultural Resources: Situation and Outlook Report*. Washington DC, 1987, 1988.
- White, H. "Instrumental Variables Regression With Independent Observations." *Econometrica* 50(1982):483-99.
- Wiens, T. E. "Peasant Risk Aversion and Allocative Behavior: A Quadratic Programming Experiment." *Amer. J. Agr. Econ.* 58(1976):629-35.

Construction of True Cost of Food Indexes from Estimated Engel Curves

William Noel Blisard and James R. Blaylock

This paper applies a technique to derive true cost-of-living indexes from the estimation of a simple system of piglog Engel curves. This technique allows a maximum amount of commodity disaggregation, and the indexes show the impact of inflation on households with different expenditure levels and demographic characteristics. Results indicate that the consumer price index (CPI) underestimated the cost of food over the 1980–85 period. The CPI more accurately reflects the cost of food for nonwhite households with low or average food expenditures.

Key words: Engel curves, true cost-of-living index.

Economists and government agencies have utilized various fixed-weight price indexes to ascertain how changing price levels affect consumers and to adjust the benefit levels of welfare and transfer programs. However, fixed weight indexes, such as the consumer price index (CPI), ignore the fact that consumers will substitute among goods as relative prices change, thereby altering the appropriate true weights. The result is that fixed weight indexes are biased, yet often used, estimators of the cost of living.

In order to construct index numbers based upon economic theory and to incorporate substitution effects by allowing weights to change from period to period, economists have developed "true indexes," which are typically derived from the estimated parameters of a complete demand system. However, demand systems tend to be limited to several broad categories of goods because of estimation problems; as such, they do not capture the substitution effects within the individual categories. Consequently, a true index that does not require the estimation of a demand system has been sought. One candidate, advanced by Diewert, and Fry and Pashardes, is the Tornqvist price index which, under specific conditions, is a true index. This index is easy to derive since it simply requires a knowledge of budget shares and prices over the relevant time

period. Unfortunately, this index may also fail to capture substitution effects from relative price changes because budget shares are fairly constant over time. This occurs if consumers make quantity adjustments as relative prices change but leave real expenditure levels almost constant.

The purpose of this paper is to derive true cost-of-food indexes for various demographic groups in the United States from estimated piglog Engel curves. In addition, we wish to determine if the CPI for total food over- or underestimates the cost-of-food for any demographic group. The indexes constructed in the study are based upon the premise that substitution effects can be captured by estimating Engel curves in which the intercepts may shift from one time period to another. These true indexes are closely related to the Tornqvist index but use the estimated intercepts from the Engel curves rather than observed budget shares as weights. The following sections outline the estimation of a true cost-of-living index from piglog Engel curves, discuss the entry of demographic variables into the model, and present the estimated Engel curves and several true cost-of-food indexes for various demographic profiles.

Estimating the True Cost Index of the Piglog Model

Piglog (price independent generalized logarithmic) models represent a specific class of pref-

William Noel Blisard and James R. Blaylock are agricultural economists with the Economic Research Service, U.S. Department of Agriculture.

The authors would like to thank two anonymous reviewers for their constructive comments.

ferences which permit exact aggregation over consumers (Muellbauer 1975, 1976). That is, the piglog functional forms represent market demands as if they were the outcome of a representative consumer. These preferences are represented by a cost or expenditure function which defines the minimum expenditure necessary to attain a specific utility level at given prices. This cost function is denoted as $c(u, p)$ for utility level u and price vector p . Thus, the piglog function is defined as

$$(1) \quad \ln c(u, p) = (1 - u) \ln a(p) + u \ln b(p),$$

where u lies between 0 (subsistence) and 1 (bliss). Thus, the positive linearly homogenous functions $a(p)$ and $b(p)$ are regarded as the costs of subsistence and bliss, respectively. Specific functional forms are assigned to $a(p)$ and $b(p)$ so that their first- and second-order derivatives can approximate any arbitrary cost function.

Within the context of the piglog model the true cost index for any household may be written as

$$(2) \quad \ln P(p_1, p_0; u_{hr}) \\ = [a(p_1) - a(p_0)] + [b(p_1) - b(p_0)]u_{hr}$$

for price vectors p_1 and p_0 and reference utility u_{hr} . Again, this expression can be interpreted as the cost of living at some minimum level of consumer expenditure, say, $\ln S_t = a(p_1) - a(p_0)$ and a marginal expenditure index, $\ln M_t = [b(p_1) - b(p_0)]u_{hr}$. Fry and Pashardes note that, over time, $\ln S_t$ incorporates the effects of substitution among goods, while differences in $\ln M_t$ across households reflect the distributional effects of inflation.

Utilizing the indirect utility function, the Marshallian budget shares of the piglog model can be derived from the above cost function as

$$(3) \quad w_{iht} = a_i(p_t) + [b_i(p_t)/b(p_t)][\ln x_{ht} - a(p_t)].$$

This complete demand system could be estimated, but the number of commodities or groups that could be considered is constraining. In general, a high degree of aggregation results in little substitution between the groups because most of the substitution occurs within the separate groupings. However, Fry and Pashardes propose modeling the substitution effects as shifts in the $a(p)$ part of the piglog cost function over time.

Specifically, when the piglog cost function takes the almost ideal demand system form, the Engel curve can be written as

$$(4) \quad w_{iht} = A_{it} + B_i [\ln x_{ht} - \alpha_i],$$

where $A_{it} = A_{i0} + \sum_j \lambda_{ij} \ln(p_{jt}/p_{j0})$, $t = 0 \dots T$, and α_i is the log of the observed minimum household expenditure. The A_{it} terms thus reflect the substitution effects as prices change from p_{j0} . Engel curves can then be estimated for a large number of commodities.

The estimated parameters of the above Engel curves are used to construct a base period referenced index series for any given household h :

$$(5) \quad \ln I_{ht} = \sum_i A_{i0} \ln(p_{it}/p_{i0}) + \sum_i A_{it} \ln(p_{it}/p_{i0}) \\ + [\Pi_i p_{it}^{B_i} - 1][\ln x_{ht} - \alpha_i].$$

The average of the first two indexes is the "reference household's" (minimum expenditure) true cost index. All other indexes are relative to the reference household's index; these indexes differ by the effect of their level of expenditure, which is the third term of equation (5).

The Tornqvist index is formally defined as

$$.5(w_{iht1} + w_{iht0}) \ln(p_{it}/p_{i0}),$$

where w_{iht} is the budget share for the i th good for the h th household in period $t = 0, 1$. However, instead of observed budget shares, we substitute the estimated intercept terms from the piglog Engel curves. These estimated intercepts represent the budget shares of the "reference household" and capture the substitution effects that occur as prices change from a base period. In addition, we can derive indexes for households with expenditures above that of the reference household by utilizing the marginal expenditure index.

The Tornqvist index is a true cost-of-living index if the underlying cost function is translog (Diewert) or quadratic (Fry and Pashardes). Since Engel curves related to the almost ideal demand system have an underlying quadratic cost function, the derived indexes are considered true indexes.

Incorporating Demographics into the Model

Household characteristics affect patterns of demand and cause price changes to have varying effects on the cost of living for different households. For illustration consider a household characteristic, say z , which is a continuous variable. The cost function may be written as

$$(6) \quad \ln c(u_h, p, z_h) \\ = a(p) + b(p)u_h + d(p) \ln z_h,$$

where $a(p)$ and $b(p)$ are defined above and $d(p) = \epsilon + \sum_{i=1}^r \ln p_{it}$.

Intercept shifts for each time period capture the substitution effects, and the Engel curves are estimated as

$$(7) \quad w_{iht} = A_{it} + B_{it}(\ln x_{it} - \alpha_t - \eta \ln z_h) + \zeta_i \ln z_h,$$

where η is the equivalent income scale at base period prices (Fry and Pashardes). Our strategy was to group the data by the z demographic variable, so that the $\eta \ln z_h$ term could be absorbed into the definition of the minimum household expenditure, α_t , for the appropriate demographic group.

Traditional zero-one dummy variables can account for various noncontinuous demographic effects such as race and region for both the intercept and the slope parameters. Demographic effects for race, region, and household size are captured in this study.

Empirical Results

True cost-of-food indexes were constructed from Engel curves estimated from data taken from the "Continuing Consumer Expenditure Survey" (CCES) for the years 1980 through 1985. The CCES contains two components, each with its own questionnaire and sample. The first is an interview panel survey in which approximately 5,000 households are surveyed every three months over a one-year period. The second is a diary survey of approximately the same sample size in which households keep an expenditure diary survey for two consecutive weeks. This latter survey obtains data on small, frequently purchased items that are normally difficult to recall, including food and beverages.

By using this survey we considered sixteen food categories: beef, cereal and bakery products, dairy products, eggs, food-away-from-home, fresh fruit, fish, fats and oils, fresh vegetables, nonalcoholic beverages, other meats, pork, processed fruit, processed vegetables, poultry, and sugar and sweeteners. The individual CPI's which correspond to the above sixteen food groups and which comprise the CPI for total food were used as proxies for price. The estimated equation for each of the sixteen food groups is

$$(8) \quad W_{iht} = A_{it} + A_{inc}D_{nc} + A_{is}D_s + A_{iw}D_w + A_{ir}D_r + Z_{iz} \ln z_h + (Y_{it} + Y_{inc}D_{nc} + Y_{is}D_s + Y_{iw}D_w + Y_{ir}D_r)(\ln x_{it} - \alpha_t),$$

where $t = 1980 \dots 1985$ and the A and Y sub-

scripted variables are dummy shifters for the intercept, A_{it} , and slope, Y_{it} , for the North Central, South, West, and race. In addition, we have the intercept shift parameter for household size, Z_{iz} . For this variable z_h is the log of the family size equivalent scales implicit in the official poverty thresholds published by the Bureau of the Census. $\ln X_{ht}$ is logged household expenditure on total food. Finally, α_t is the log of the observed minimum household expenditure on total food for each demographic group, e.g., nonwhite households in the Northeast or nonwhite households in the South.

The estimates for the sixteen Engel curves are presented in table 1. For each equation, A_{80} through A_{85} represents the intercept for the Northeast for each year of data. A_{nc} through A_w represent regional dummy variables for the North Central, South, and West. Z is the estimated coefficient for household size, and A_r is the demographic dummy variable for race.

Slope expenditure parameters are represented by Y through Y_r , where Y represents the estimated expenditure coefficient for nonwhites in the northeast, and Y_{nc} , Y_s , and Y_w , are the estimated dummy slope shifters for expenditures by nonwhites in the North Central, South, and West, respectively. Y_r is the dummy expenditure slope shifter for whites. R^2 is the goodness of fit of each equation, and F is a significance test between estimating an intercept for each year versus one common intercept for all years. Many of the estimated coefficients are highly significant.

Variation in the intercepts is a necessary condition the presence of substitution effects and the F -tests indicate that most equations are better represented by allowing the intercept to shift from one period to another versus a single estimated parameter. Exceptions include fresh fruit, fish, processed vegetables, and sugar and sweeteners. Apparently, little substitution occurs between these four categories and the others. For instance, households may substitute one kind of fruit for another but may not substitute fresh vegetables for fresh fruit. In addition, the Pearson correlation coefficient between changes in the intercept terms and changes in the individual CPIs was .73, which supports the argument that shifts in the intercept terms capture the substitution effects. A rigorous proof of this would be to calculate the implied substitution coefficients and then check for concavity. However, we have more parameters than equations and cannot provide this check. We will return to the subject of substitution below.

Comparing the estimated intercepts for any

Table 1. Parameter Estimates of Engel Curves

Commodity	A80	A81	A82	A83	A84	A85	A _{nc}	A _s
Beef	.0748 (.003)	.0722 (.003)	.0657 (.003)	.0645 (.003)	.0610 (.003)	.0541 (.003)	-.0084 (.002)	-.0053 (.002)
Cereal & bakery	.1199 (.002)	.1225 (.003)	.1224 (.003)	.1217 (.003)	.1233 (.003)	.1272 (.003)	-.0021 (.002)	-.0029 (.002)
Dairy	.0888 (.003)	.0905 (.003)	.0882 (.003)	.0889 (.003)	.0839 (.003)	.0869 (.003)	.0023 (.002)	-.0089 (.002)
Eggs	.0228 (.001)	.0230 (.001)	.0221 (.001)	.0216 (.001)	.0214 (.001)	.0191 (.001)	-.0004 (.001)	-.0009 (.001)
Food-away- from-home	.2524 (.009)	.2495 (.009)	.2622 (.009)	.2667 (.008)	.2757 (.008)	.2763 (.008)	.0353 (.007)	.0563 (.007)
Fresh fruit	.0414 (.002)	.0422 (.002)	.0447 (.002)	.0428 (.002)	.0429 (.002)	.0422 (.002)	-.0029 (.001)	-.0024 (.001)
Fish	.0330 (.001)	.0326 (.001)	.0327 (.001)	.0326 (.001)	.0323 (.001)	.0338 (.001)	-.0073 (.001)	-.0049 (.001)
Fats and oils	.0237 (.001)	.2336 (.001)	.0221 (.001)	.0206 (.001)	.0217 (.001)	.0225 (.001)	.0007 (.009)	-.0008 (.001)
Fresh vegetables	.0392 (.002)	.0430 (.001)	.0435 (.001)	.0438 (.001)	.0437 (.001)	.0427 (.001)	-.0027 (.001)	-.0013 (.001)
Non- alcoholic beverage	.0732 (.003)	.0679 (.003)	.0678 (.002)	.0709 (.002)	.0714 (.002)	.0721 (.002)	-.0054 (.002)	-.0056 (.002)
Other meats	.0345 (.002)	.0364 (.002)	.0360 (.002)	.0351 (.001)	.0332 (.001)	.0341 (.002)	-.0032 (.001)	-.0094 (.001)
Pork	.0498 (.002)	.0518 (.002)	.0478 (.002)	.0464 (.002)	.0447 (.002)	.0450 (.002)	.0058 (.002)	.0064 (.002)
Processed fruit	.0405 (.001)	.0414 (.001)	.0419 (.001)	.0424 (.001)	.0413 (.001)	.0433 (.001)	-.0061 (.001)	-.0075 (.001)
Poultry	.0558 (.002)	.0551 (.002)	.0051 (.002)	.0530 (.001)	.0553 (.002)	.0530 (.002)	-.0092 (.001)	-.0040 (.001)
Processed vegetables	.0218 (.001)	.0222 (.001)	.0222 (.001)	.0026 (.001)	.0028 (.001)	.0216 (.001)	-.0016 (.001)	-.0003 (.001)
Sugar and sweeteners	.0284 (.002)	.0264 (.002)	.0254 (.002)	.0263 (.002)	.0253 (.002)	.0262 (.002)	.0041 (.001)	.0015 (.001)

^a Three asterisks indicate significant at .01; double asterisk indicates significant at .10 level; single asterisk indicates significant at .05 level.

given year with the preceding year indicates the food group substitution pattern for the representative household. In comparing 1981 with 1980 we see that the budget shares for beef, food-away-from-home, fish, fats and oils, nonalcoholic beverages, and poultry decreased, while the remaining categories increased. Hence, the representative households substituted away from these categories into the others.

As noted above, true cost indexes for reference households can be calculated from the estimated intercepts of the Engel curves. Marginal "demographic" indexes, calculated from the coefficients that shift the intercepts, can be utilized to construct indexes which take into account the effects of race, region, and household size. In turn, marginal "expenditure" indexes, calculated from coefficients which shift the slopes of the Engel curves, can be utilized to construct indexes which take into account expenditures

above those of the reference household by race and region.

A true cost-of-food index was constructed for a reference household defined as a nonwhite single-person household in the Northeast; it is presented in table 2. In addition, we have indicated how the reference household can be adjusted to account for demographic effects by race, region, and household size, as well as marginal expenditure effects by race and region. Over the 1980-85 period the true cost of food for the reference household rose 21.8%. White reference households experienced a greater inflation rate than the nonwhite reference household. These values ranged from .1% higher in 1981 to .7% in 1983. Similarly, the three regional marginal indexes are all greater than 100.0, which indicates that the reference household in the Northeast experienced the lowest rate of price increase. While both the North Central and South

Table 1. Continued.

A_w	Z	A_r	Y	Y_{nc}	Y_s	Y_w	Y_r	R^2	F
-.0076 (.002)	.0472 (.002)	-.0080 (.002)	.0119 (.002)	.0052 (.002)	.0035 (.002)	-.0031 (.002)	-.0054 (.001)	.44	26.23***
-.0201 (.002)	.0184 (.002)	-.0018 (.002)	-.0315 (.002)	-.0074 (.002)	-.0020 (.002)	.0069 (.002)	.0085 (.001)	.66	17.60***
-.0046 (.002)	.0188 (.002)	.0401 (.002)	-.0263 (.002)	-.0076 (.002)	.0025 (.002)	.0064 (.002)	-.0107 (.002)	.63	2.87**
.0007 (.001)	.0004 (.001)	-.0025 (.001)	-.0047 (.001)	-.0020 (.001)	.0009 (.001)	-.0015 (.001)	-.0010 (.001)	.30	11.25***
.0637 (.008)	-.1140 (.006)	.0026 (.006)	.0608 (.007)	.0085 (.007)	-.0060 (.007)	-.0062 (.007)	.0227 (.006)	.66	6.62***
.0072 (.001)	-.0111 (.001)	-.0011 (.001)	-.0014 (.001)	-.0005 (.001)	-.0022 (.001)	-.0059 (.001)	-.0025 (.001)	.36	1.71
-.0042 (.001)	.0017 (.001)	-.0135 (.001)	.0064 (.001)	-.0014 (.001)	.0008 (.001)	.0005 (.001)	-.0024 (.001)	.21	.54
.0022 (.001)	.0040 (.008)	.0037 (.008)	.0010 (.001)	-.0010 (.001)	-.0011 (.001)	-.0033 (.001)	-.0027 (.001)	.38	4.34***
.0082 (.001)	-.0047 (.001)	-.0066 (.001)	.0010 (.001)	-.0015 (.001)	-.0013 (.001)	-.0064 (.001)	-.0001 (.001)	.43	5.00***
-.0091 (.002)	.0018 (.002)	.0145 (.002)	-.0183 (.002)	.0080 (.002)	.0081 (.002)	.0081 (.002)	-.0026 (.001)	.48	3.16***
-.0134 (.001)	.0124 (.001)	-.0015 (.001)	.0006 (.001)	.0009 (.001)	.0002 (.001)	.0013 (.001)	-.0027 (.001)	.33	2.79**
-.0035 (.002)	.0207 (.002)	-.0174 (.001)	.0109 (.002)	-.0015 (.002)	-.0054 (.002)	-.0028 (.002)	-.0065 (.001)	.35	6.95***
-.0058 (.001)	-.0055 (.001)	-.0048 (.001)	-.0049 (.001)	.0012 (.001)	.0027 (.001)	.0023 (.001)	.0003 (.001)	.34	1.88*
-.0102 (.001)	.0063 (.001)	-.0215 (.001)	-.0019 (.001)	-.0003 (.001)	-.0005 (.001)	.0012 (.001)	.0008 (.001)	.31	2.19*
-.0027 (.001)	.0045 (.001)	-.0029 (.001)	-.0033 (.001)	.0020 (.001)	.0021 (.001)	.0012 (.001)	.0022 (.001)	.32	.85
-.0008 (.001)	.0038 (.001)	.0006 (.001)	-.0033 (.001)	-.0025 (.001)	-.0022 (.001)	.0012 (.001)	.004 (.001)	.28	1.70

had similar rates of price increases, the West experienced the highest rates of increase.

The demographic marginal indexes for household size are included in table 2. Each value for household size 2 through 5 is below 100.0 after 1980, generally decreasing in magnitude over the six years. This pattern indicates that, relative to a single-person household, the true cost-of-food falls as household size increases. Intuitively, this may seem contradictory. However, the sixteen food categories include food-away-from-home, which experienced one of the largest price increases of all food categories. The data indicate that per capita food spending declines for this category as household size increases. In 1981 a two-person household experienced an inflation rate that was .1% lower than a single-person household, while a five-person household experienced a rate that was .4% lower than a single-person household. In 1985 the cumulative rates experienced by a two- and five-person household were .5% and 1.8% lower, respectively, than a single-person household.

Finally, marginal expenditure indexes are also shown in table 2. These marginal indexes are used to construct true cost-of-food indexes for households with expenditures greater than the reference household. They indicate the response of the reference index to a 1% increase in total food expenditure. The race variable for white households is again greater than 100.0, indicating that the true cost-of-food index increases as expenditures increase above that of the reference household relative to nonwhites. However, the three regional expenditure indexes are all less than 100.0, which indicates that consumers in the Northeast have a larger marginal expenditure index than consumers in the other three regions. These regions have similar marginal expenditure indexes and are just slightly less than that of the Northeast. Hence, while the Northeast should have the lowest value index for the reference household, those Northeast households with expenditures greater than the reference household may have true indexes greater than those of the North Central, South, and West.

Table 2. Reference Household Index and Effects of Demographic Adjustments

Year	Reference Index	Marginal Demographic Region				Household Size				Marginal Expenditure Region			
		Northeast Race	North		West	2	3	4	5	Race	North		West
			Central	South							Central	South	
1980	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
1981	107.9	100.1	100.1	100.1	100.2	99.9	99.8	99.7	99.6	100.0	99.9	99.9	99.9
1982	112.0	100.5	100.3	100.3	100.4	99.8	99.7	99.5	99.4	100.1	99.9	99.8	99.9
1983	114.2	100.7	100.4	100.5	100.6	99.7	99.5	99.1	98.9	100.3	99.9	99.8	99.9
1984	119.0	100.5	100.5	100.4	100.7	99.7	99.3	99.0	98.7	100.3	99.9	99.8	99.8
1985	121.8	100.2	100.4	100.7	100.9	99.5	99.1	98.6	98.2	100.4	99.9	99.7	99.8

Table 3 elaborates upon the effects of substitution obtained from using the intercept parameters to construct relevant indexes. The first column repeats the reference index from table 2. The second column is the fixed base expenditure weighted index. This index lies below the Tornqvist index because these effects refer to substitution between expenditures rather than quantities of goods. The current weighted index lies above the Tornqvist as expected. These results suggest that an index which uses expenditure shares as weights can be a good approximation to the true index in the case of total food. Conversely, this implies that a quantity share index would be a poor approximation to the true index.

A quantity-share-weighted index is calculated using the fact that $A_{it} = p_{it}q_{it}/x_t$, where q_{it} is the quantity purchased by the reference household. Since the Tornqvist index is the true cost-of-food index corresponding to x_t , $A_{it}S_{it}/p_{it} = \omega_{it}$ is the quantity share of the i th good. Thus, $\sum \omega_{i0} \ln(p_{it}/p_{i0})$ and $\sum x_{it} \ln(p_{it}/p_{i0})$ are the fixed based and current weighted indexes which use the quantity shares as weights. These indexes are reported in the last two columns of table 3.

As expected, the index which uses the base period quantity shares overstates the cost of food, while the index which uses the current period quantity shares understates the true cost of food. Importantly, the biases of these two indexes are larger than those of the expenditure indexes. Thus, the shifts of the intercept terms of the Engel curves appear to capture the substitution effects resulting from relative price changes.

With this background, other cost-of-food indexes constructed from the estimated Engel curves can be considered. Table 4 indicates constructed indexes for a reference (least expenditure) single-person household as well as for single households with average and high expenditures.¹ Indexes were constructed for the total sample (all singles), nonwhites and whites, as well as by region. Average expenditure (the average weekly household expenditure in the sample) ranged from \$56.87 in 1981 to \$67.60 in 1985. The high expenditure level was one standard deviation above the mean values. The standard deviation was approximately \$40.00 for each year in the sample.

¹ Conditional expenditure elasticities can be calculated from a knowledge of budget shares and the expenditure coefficients using the formula $\epsilon_i = 1 + B_i/w_i$. Contact the authors for the appropriate budget shares.

Table 3. True Index for Reference Household and Substitution Effects

Year	Reference Index	Expenditure Share Base	Weighted Current	Quantity Share Base	Weighted Current
1980	100.0	100.0	100.0	100.0	100.0
1981	107.9	107.2	108.6	108.7	107.1
1982	112.0	111.0	113.0	113.1	110.9
1983	114.2	112.7	115.7	115.8	113.5
1984	119.0	117.7	120.3	121.0	117.3
1985	121.8	119.0	124.6	124.8	118.5

Table 4 and the least expenditure indexes indicate that each individual index is above the CPI. This relationship also holds for those households with expenditures above the reference household. Other things equal, the true index should lie below the CPI since it allows for substitution among the sixteen food categories. However, over the 1980–85 period households increased their budget share of food away from home. This category also had a large price increase over the sample period. However, because the CPI is a fixed weight index, it underestimates the increase in total food prices. Note also that the indexes in table 4 are for single-person households, and that the data indicate that these households allocate a larger budget share to food away from home than do larger households.

In the individual categories, whites have a higher index than nonwhites and the Northeast and West have the lowest and highest indexes of the four regions, respectively. This result was expected from the discussion of the demographic marginal indexes. For the reference households the differences in the indexes for the races are slight, amounting to .1 of a point in 1981 and .2 of a point in 1985. Similarly, differences among regions are small, ranging from .2 of a point difference between the Northeast and the West in 1981 to .9 of a point between the same two regions in 1985.

When expenditures above that of the reference household are considered, whites have a higher average expenditure index than nonwhites, but now the South has the lowest index, while the West again has the highest. This pattern occurs because the South has the lowest marginal expenditure index of the four regions. In general, the difference in the indexes between the races is greater than between regions. This result is not totally unexpected because differences in income received are probably greater

between the races than between the regions. Hence, even given the same dollar amount of food expenditures, buying patterns between the races is different, as was shown by the race variable in the estimated Engel curves. While part of this effect is surely due to asset or wealth disparity, some may result from cultural differences.

Considering the high expenditure level in table 4, the same pattern holds that appeared with the average expenditure category, except the differences between the races again widen. Thus, whites have a true cost-of-food index which is .4 of a point higher than nonwhites in 1981, and 1.3 points higher than nonwhites in 1985. Undoubtedly, the difference is because whites allocate a larger share of their food expenditures to the food-away-from-home category.

Table 5 indicates the true cost-of-food index for an average size family with average food expenditures with the same demographic characteristics as table 4. As shown in the table, average family size over the period in question was 2.5 people. Relative to the CPI, the true cost-of-food index for the total sample is still greater but by a smaller margin as a result of the negative effect of household size. Hence, the true cost index is .3 of a point higher than the CPI in 1981 and 1.4 points higher in 1985. However, the index for nonwhites is much closer to the CPI. The indexes are the same in 1981 and differ by .5 of a point in 1985. Across regions all the true indexes are above the CPI; again, the West has the largest true cost-of-food index as before.

When household size is increased to four people in table 5, the same pattern is found except the calculated index for nonwhites is below that of the CPI. However, the index for the total sample is very close to the CPI, .2 for a point higher in 1981 and .5 for a point higher in 1985. The largest difference occurs in 1984 when the

Table 4. Single-Person Household Indexes

Year	CPI	All	Nonwhite	White	Northeast	North Central	South	West
Single Reference Household								
1980	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
1981	107.8	108.2	108.1	108.2	108.1	108.2	108.2	108.3
1982	112.2	112.8	112.3	112.8	112.5	112.8	112.8	112.9
1983	114.5	115.3	114.9	115.4	114.9	115.3	115.4	115.5
1984	118.9	120.0	119.5	120.1	119.6	120.1	120.0	120.4
1985	121.7	122.5	122.3	122.5	121.8	122.2	122.6	123.2
Single Average Expenditure Household								
1980	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
1981	107.8	108.3	108.0	108.4	108.3	108.3	108.2	108.4
1982	112.2	113.4	112.7	113.5	113.3	113.5	113.2	113.5
1983	114.5	116.3	115.1	116.4	116.1	116.4	116.1	116.5
1984	118.9	121.0	119.8	121.2	120.9	121.2	120.8	121.2
1985	121.7	124.0	123.0	124.1	123.6	123.4	123.7	124.6
Single High Expenditure Household								
1980	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
1981	107.8	108.3	108.0	108.4	108.4	108.3	108.2	108.4
1982	112.2	113.6	112.8	113.6	113.5	113.7	113.3	113.6
1983	114.5	116.5	115.3	116.7	116.4	116.6	116.3	116.7
1984	118.9	121.3	119.9	121.4	121.2	121.4	121.0	121.4
1985	121.7	124.4	123.2	124.5	124.0	124.3	124.0	125.0

Table 5. Comparison of CPI and the True Cost of Food for the Average Size Household and Households of Four People

	CPI	All	Nonwhite	White	Northeast	North Central	South	West
Household size = 2.5								
1980	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
1981	107.8	108.1	107.8	108.2	108.1	108.1	108.0	108.2
1982	112.2	113.1	112.4	113.2	113.0	113.2	112.9	113.2
1983	114.5	115.8	114.6	116.0	115.6	115.9	115.6	116.0
1984	118.9	120.4	119.2	120.6	120.2	120.6	120.2	120.5
1985	121.7	123.1	122.2	123.2	122.7	123.0	122.9	123.7
Household size = 4								
1980	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
1981	107.8	108.0	107.7	108.1	108.0	108.0	107.9	108.1
1982	112.2	112.8	112.1	112.9	112.7	112.9	112.6	112.9
1983	114.5	115.3	114.1	115.5	115.1	115.4	115.1	115.5
1984	118.9	119.8	118.6	120.0	119.8	120.0	119.8	119.9
1985	121.7	122.2	121.3	122.3	121.8	122.1	122.0	122.8

true index is .9 of a point higher than the CPI. The regional patterns are the same as before.

Conclusions

We have applied a technique in which true cost-of-food indexes are derived from the estimation of a simple system of Engel curves. This method demonstrates the feasibility of deriving cost-of-food indexes without estimating a complete demand system. Consequently, a large number of commodity groups can be considered and the substitution effect between the groups captured. Also, indexes for various demographic groups in society can be constructed from survey data.

In general, the results indicate that the CPI underestimated the cost of food over the 1980–85 period. The CPI more accurately reflects the cost of food for nonwhite households with low or average food expenditures and with four or more household members. Conversely, the CPI underestimates food costs the most for small white households with average or above food expenditures. However, most of the true cost-of-food indexes which we calculated were close to the CPI for total food. In this sense, the CPI was a fairly good indicator of total food costs for the above groups over the study period.

[Received May 1990; final revision received September 1990.]

References

- Diewert, W. E. "Exact and Superlative Index Numbers." *J. Econometrics* 4(1976):115–45.
- Fry, V. C., and P. Pashardes. "Constructing the True Cost-of-Living Index from the Engel Curves of the Piglog Model." *J. Appl. Econometrics* 4(1989):41–56.
- Muellbauer, J. "Aggregation, Income Distribution and Consumer Demand." *Rev. Econ. Stud.* 62(1975):525–43.
- . "Community Preferences and the Representative Consumer." *Econometrica* 44(1976):979–99.
- Tornqvist, L. "The Bank of Finland's Consumption Price Index." *Bank of Finland Monthly Bull.* 10(1936):1–8.
- U.S. Department of Labor, Bureau of Labor Statistics. *The Consumer Price Index: Concepts and Content Over the Years*. Washington DC, 1978.

Wife's Employment, Food Expenditures, and Apparent Nutrient Intake: Evidence from Canada

Susan Horton and Cathy Campbell

Women's employment has important effects on the share of food in total expenditure, the share of food-away-from-home in the food budget, cost per calorie, and per capita availability of twelve nutrients. Empirical results are presented using Canadian household food expenditure survey data. As expected, wife's employment increases the share of food-away-from-home in the food budget. Wife's full-time employment has a negative effect on apparent nutrient intake, not offset by the higher per capita income of such households. There are policy implications for nutrition education and nutrient content of food-away-from-home.

Key words: apparent nutrient intake, Canada, food-away-from-home, food expenditures, household survey, wife's employment, women's employment.

The increase in women working outside the home has had important economic effects. One effect involves food expenditures, both in terms of patterns of eating out and the types of foods prepared in the home, with consequent effects on nutrient availability. This study examines the effect of wife's employment on an extensive set of variables describing food behavior. The variables include the share of food in total expenditure, the share of food-away-from-home (FAFH) in the food budget, price paid per calorie, and the per capita availability of twelve nutrients. The results are of potential interest for food policy, for nutrition education, and for programs which aim to affect the nutrient composition of food eaten outside the home (restau-

rant food, fast food, school meals, etc.). In addition, this is the first time that public use data tapes of Canadian national food expenditure surveys have been available for use in research. The remaining sections of this paper review the literature, discuss the methods, data, results, and conclusions of the study.

Literature Review

Several empirical studies have examined the effects of wife's employment status, mostly for the United States. Expenditures on food-away-from-home generally increase where the wife works outside the home (e.g., McCracken and Brandt, Kinsey, Sexauer, and two studies cited in Morgan, by Morgan and Goungetas and by Haines). Capps, Tedford, and Havlicek also found that households where the wife is not employed outside the home spent a larger share of the food budget on nonconvenience foods and a smaller share on "complex convenience foods." Barewal showed that households in Canada headed either by a single adult or by a couple where both worked full time outside the home spent 50% more on restaurant food than households headed by a couple where the wife worked full time in the home. However, because this study used only univariate analysis, the wife's

Susan Horton is an associate professor, Department of Economics, University of Toronto; Cathy Campbell is an assistant professor, Division of Nutritional Sciences, Cornell University.

This analysis is based on the 1984 Survey of Family Expenditures, Statistics Canada. All computations on these data were done by the authors, and responsibility for their use and interpretation is entirely that of the authors.

The authors would like to thank the Family Expenditures Section of Statistics Canada and in particular Ulysse Neveu for assistance in obtaining the household data. Thanks to Agriculture Canada and in particular Linda Robbins for assistance in obtaining the nutrient composition data for specific foods. The authors are grateful to Duncan Thomas and three anonymous referees for helpful comments on an earlier draft. Paul McGuire gave invaluable assistance with some of the computations, and Monica Neitzert and Douglas Tisdall provided capable research assistance.

work status effects could not be distinguished from other household composition and income effects. Fewer studies have considered the effects of wife's employment status on nutrition, and the findings of these studies are less conclusive. Adrian and Daniel (using expenditure data) found no significant difference in nutrient intake between households where the wife was employed and those in which the wife was not employed.

For developing countries the effects of wife's employment status has been widely studied, often within the context of new household economics models. The effect on food expenditures has received little attention because restaurant expenditures are far less important than in high income countries. However, one study by Senauer, Sahn, and Alderman documented that a higher wage (actual or imputed) of the wife was associated with a shift in consumption away from rice (which involves more preparation time) and toward bread (which is more convenient) in Sri Lanka. Several studies have used anthropometric measures to examine the effect of wife's employment on nutritional status. Tucker and Sanjur surveyed a number of studies for the Philippines, Ghana, and Kerala and found mixed effects. In Tucker and Sanjur's study of Panama, wife's employment status was positively associated with intake of protein, calories and calcium, and with hemoglobin level and anthropometric measures.

Methods

Women's employment outside the home has several possible effects on consumption patterns. First, higher household income from the wife's employment would increase consumption of all normal goods and increase the budget share of luxury goods. Second, the new household economics framework suggests that wife's employment would cause a shift from consumption of time-intensive toward goods-intensive commodities. Other effects result from changes in control over household income. Income earned by the wife may give women a greater say in household expenditure decisions, especially for expenditures on children and on basic needs including food (e.g., Dwyer).

The present study uses three measures of food expenditures: the share of income spent on food, the share of the food budget allocated to food-away-from-home, and cost per calorie of home-prepared food (food purchased in food stores).

The last variable reflects the widely observed phenomenon that as incomes rise, households allot more of the food budget to changing the type, quality, and variety of foods rather than increasing the quantity of calories consumed. The study also examines daily per capita availability of twelve nutrients (calories, protein, fat, carbohydrate, folate, calcium, iron, vitamin A, vitamin C, thiamine, riboflavin, and niacin).

The following budget equations are estimated:

$$\begin{aligned} \text{food share of income} &= f(Y, P, Z), \text{ and} \\ \text{FAFH share of food expenditure} &= g(Y, P, Z), \end{aligned}$$

where Y is per capita income, P is a vector of prices, and Z is a vector of demographic characteristics including wife's employment status.¹ Nutrient availability per capita and cost per calorie (an indicator of tastes for noncaloric aspects of food) are treated as derived demands, and similar equations are estimated:²

$$\begin{aligned} \text{cost per calorie} &= h(Y, P, Z), \text{ and} \\ \text{nutrient availability} &= n(Y, P, Z). \end{aligned}$$

A quadratic specification of income per capita was used (other specifications such as log-linear and log-quadratic were also tried, with similar results). Since cross-section data are used, region and seasonal dummies capture price effects. Household composition is represented by the number of adults sixteen to sixty-five years, the number of adults over sixty-five, and the number of children. These variables control both for scale effects as well as differences in nutrient requirements by age. Entering household composition variables directly is more flexible than using adult equivalency scales. Unfortunately, the data do not indicate the sex composition of the household (other than information on headship). Dummy variables for education of the head of household were included (some post-secondary other than a university degree/university degree/with the omitted dummy being secondary education or less). The male household head's education is the one which is used in households headed by couples. Dummy variables are also used to capture wife's working status (households with a male head, no wife/

¹ Per capita income equals total household income divided by number of household members.

² Per capita nutrient availability equals total nutrients available per household divided by number of household members. To allow for the fact that individual requirements vary by age, age composition variables were entered separately (the data tape did not contain details on sex composition of the households).

households with a female head, no husband/ households headed by a couple where the wife works full time in the market/ households headed by a couple where the wife works part time in the market/ with the omitted dummy being households headed by a couple where the wife works full time in the home).³ A house ownership dummy was included (the omitted dummy being renters and tenants) because house ownership frequently affects the facilities available for home food preparation and lifestyle differences.

Because the wife's employment variable is considered endogenous, instrumental variables were used for the two dummy variables reflecting wife's work status (full-time and part-time market work).⁴ These instruments were the following: wife's country of birth (North America, the UK, other European countries, and the rest of the world), wife's first language (English, French, or other), and wife's age (6 variables representing 10-year age groups).⁵

The estimation method was ordinary least squares (OLS), with instrumental variables for the wife's work status. For the FAFH share variable, 20.3% of the sample did not eat out in the reference period. Zero observations cause potential estimation problems. If these are true corner solutions, tobit is the preferred approach for single equation estimates. Wales and Woodland, and Lee and Pitt make suggestions as to an appropriate estimation technique where a set of budget equations is estimated. However, if the zeroes result from infrequent purchase and/ or measurement error, then OLS is the most appropriate method (Deaton, Ruiz-Castillo and Thomas). It is likely that almost all households purchase some food-away-from-home at some point during the year; therefore, OLS is the appropriate technique. Because some previous studies (e.g., McCracken and Brandt, Kinsey) have chosen to use tobit when analyzing food-away-from-home, tobit estimates for this equation are presented in an appendix table, purely for comparative purposes. OLS was also used

in the equations using instruments to predict values for the wife's work choice variables.

Given this model, the likely effect of wife's work status on the food variables of interest now can be predicted. The equations control for the effect via household income. Therefore, one observed effect is caused by differences in the value of time (wife's employment would cause a substitution away from consumption of time-intensive commodities and toward goods-intensive commodities). The other effect results from wife's bargaining power in expenditure decisions. The effect on the share of food in total expenditures is ambiguous because it is not clear whether food is a time- or goods-intensive commodity, and whether wives with greater bargaining power would necessarily buy more food. Wife's market work likely would increase the share of food-away-from-home in the food budget, as well as the cost per calorie of home produced food. The effect on nutrient intake is not clear a priori.

Data

The data are from the 1984 Family Food Expenditure Survey of Canada. These data have been collected almost every two years since 1953 (Statistics Canada). Previous published studies of the Canadian Family Food Expenditure Survey have been restricted to cross-tabulations of the data (Barewal, Hunt, Karamchandani and Robbins, Mathieson and Robichon-Hunt, Nevraumont, Robbins).

In the 1984 survey, information was collected on socioeconomic characteristics of the household, and a diary on type of food, quantity, and expenditure by type of store for detailed food categories was maintained for two consecutive weeks. Food expenditures in restaurants were recorded separately, by type of restaurant and type of meal (breakfast, lunch, dinner, snack). Of the original data tape covering 5,542 households in fifteen cities, 5,188 records were usable for the expenditure analysis, and 4,746 for the nutrient availability analysis. The small sample loss included households who only participated in the first week of the survey (less than 1% of households), households who did not state their income (about 4% of the sample), and households who reported either zero income or zero food expenditure (less than 1% of the sample). Variable means for the excluded households were examined separately and did not differ significantly from the rest of the sample in

³ Statistics Canada defines full-time work as an individual working more than 48 weeks in the last 12 months, of which 25 weeks must be full-time employment. Part-time work is between 1 and 48 weeks, or more than 48 weeks but with less than 25 of these weeks full time.

⁴ The part of income which depends on wife's earnings is also endogenous. However, income was not given by source in the data set, and no suitable instruments for income were available.

⁵ Wife's work status = $f(\text{country of birth, first language, age})$. The instruments include variables likely to affect women's labor force participation but not highly correlated with food expenditures.

those socioeconomic characteristics of interest in this study.

The nutrient analysis was restricted to "housekeeping households," defined as households which ate at least one meal at home each week of the survey and which ate over half their meals from home supplies during the two-week period. For households eating the majority of their meals outside the home, food consumed at home provides an unreliable guide to nutrient availability. This condition excluded a further 8% of the households.

Because no recent national nutrient intake survey data for Canada are available, apparent nutrient intake was also constructed from the expenditure data.⁶ The conversion from food quantities to nutrients was based on Canadian Nutrient File conversion factors obtained from Agriculture Canada. The adjustment from nutrient availability from home sources, to total nutrient availability, was undertaken as follows. The number of meals consumed from home supplies was calculated (equal to the number of members at home times 21 meals per week, less

the number of meals consumed by guests in the household, less the number of meals consumed by household members elsewhere as guests or on expenses, less the number of meals consumed in restaurants). The ratio (M) of total meals consumed per week by household members to total meals consumed per week from home supplies was then calculated. Total nutrient availability for the household was obtained by multiplying nutrient availability from home supplies by the ratio M . Thus, we are assuming implicitly (following Adrian and Daniel, Mathieson and Robichon-Hunt, etc.) that a meal eaten away from home has the same nutrient value as the one missed at home.

However, unlike Adrian and Daniel, and Mathieson and Robichon-Hunt, we allowed for different sized meals at different times of day. The distribution of nutrients over the day was dinner 50% lunch 27%, and breakfast 23% (Kennedy et al.).⁷ Meals eaten by non-family members, and meals eaten out by household members as guests and on business expense accounts were evaluated as 33.3% of the day's nutrients. Food purchased and eaten outside the home as snacks

⁶ More accurate techniques for obtaining information on nutrient intake might be utilized; however, expenditure data provide some useful information on patterns of intake, given a large enough sample. For example, Nelson, Dyson and Paul compared nutrient data from household expenditure data and from a partially weighed method. They found no consistent bias from using one or other technique.

⁷ For example, in a household consisting of a single adult, if 2 dinners and 1 lunch were consumed in restaurants in one week, then $(7 - 2 \times 0.5 - 1 \times 0.27) = 5.75$ "day-equivalents" were consumed from home stores. Weekly intake was then $7/5.73$ times the intake from home sources.

Table 1. Means of Food Variables, by Wife Working Categories

Variable	Type of Household, by Wife's Work Status				Female Headed
	No Wife	Wife Works Full Time in Market	Wife Works Part Time in Market	Wife Works Full Time at Home	
Per capita income (\$'000/year)	16.987	16.419	11.064	9.626	11.900
Food share	0.180	0.124	0.168	0.186	0.205
FAFH share	0.341	0.288	0.234	0.164	0.225
Cost per '00 calories(¢)	0.182	0.183	0.167	0.161	0.179
Apparent per capita daily nutrient intake					
Calories (Kcal)	2,924.1	2,329.6	2,316.4	2,363.7	2,469.0
Protein (g)	103.0	83.8	81.8	81.6	83.6
Fat (g)	134.0	104.0	101.1	104.3	107.4
Carbohydrate (g)	333.9	269.9	274.6	279.9	299.0
Folate (micrograms)	240.4	191.3	184.6	185.0	208.0
Calcium (mg)	1,220.1	960.6	944.5	948.5	1,032.6
Vitamin A (international units)	9,174.7	7,705.6	7,370.9	7,667.2	9,082.2
Vitamin C (mg)	106.0	92.7	83.5	85.7	101.4
Iron (mg)	17.9	14.1	14.0	14.3	15.0
Thiamine (mg)	1.89	1.52	1.55	1.57	1.63
Riboflavin (mg)	2.38	1.94	1.90	1.91	2.09
Niacin (niacin equivalents)	46.4	38.0	36.8	37.1	38.4
Sample size	513	1,010	839	1,271	1,113

Notes: Means given to at least 3 significant figures. Canadian dollars are used throughout the study. See variable definitions in table 2.

(1.5% of total food expenditure) was not considered.

Three deficiencies in the data should be cited. For certain variables (particularly age and income) extreme values were not coded exactly, but as a range (e.g., age over 75 was coded as "76 and over"), in order to protect confidentiality.⁸ Second, the absence of information on sex composition of the household was a problem for the nutrient analysis. Third, Statistics Canada food item categories do not correspond exactly to Agriculture Canada Canadian Nu-

trient File categories. For example, Statistics Canada combines expenditures on lemons and limes, whereas Agriculture Canada has separate nutrient conversion factors for these two items.⁹

Results

Table 1 contains results from univariate analysis of the data. Mean values for some food and nutrient variables are provided for five groups of

⁸ A point estimate was used for these variables.

⁹ Agriculture Canada provided data on the average proportion of expenditure allotted to individual items within these mixed categories to permit conversion to nutrients.

Table 2. Variable Definitions, Means, and Standard Deviations, for Regressions

Variable	Mean	Standard Deviation
Per capita income (annual: in \$'000)	13.135	9.331
Per capital income squared	259.591	589.384
Number of adults aged 16-65	1.807	1.108
Number of adults aged 65 and over	0.236	0.535
Number of children	0.630	0.978
Proportion house owners	0.570	0.495
Proportion with heads with secondary education only (omitted dummy)	0.608	0.488
Proportion with heads with some post-secondary education	0.246	0.431
Proportion with heads with university degree	0.146	0.353
Proportion female head, no husband	0.242	0.428
Proportion male head, no wife	0.138	0.345
Proportion where wife works full time in market	0.191	0.393
Proportion where wife works part time in market	0.176	0.381
Proportion where wife works full time in home (omitted dummy)	0.253	0.435
Quarter 1 (omitted dummy)	0.254	0.435
Quarter 2	0.251	0.433
Quarter 3	0.246	0.434
Quarter 4	0.249	0.435
Maritimes	0.167	0.373
Quebec	0.165	0.371
Ontario (omitted dummy)	0.233	0.423
West (Prairies and British Columbia)	0.435	0.496
Food share (annual expenditure/annual income)	0.171	0.198
FAFH share (FAFH expenditure over two weeks/food expenditure over two weeks)	0.256	0.242
Cost per '00 calories for home-prepared foods (¢)	0.173	0.073
Sample size		5188
Apparent nutrient intake (per capita day)		
Calories (Kcal)	(2,700) 2,433.326	1,572.935
Protein (g)	(56) 84.882	59.462
Fat (g)	(-) 107.609	80.898
Carbohydrate (g)	(-) 287.132	201.980
Folate (microg)	(200) 197.642	127.558
Calcium (mg)	(800) 999.446	622.295
Vitamin A (international units)	(1,000) 8,117.795	9,706.513
Vitamin C (mg)	(30) 92.683	86.425
Iron (mg)	(10) 14.754	10.210
Thiamine (mg)	(1.4) 1.607	1.159
Riboflavin (mg)	(1.7) 2.008	1.341
Niacin (niacin equivalents)	(18) 38.543	26.658
Sample size		4,746

Note: Figures in brackets are daily recommended nutrient intakes for males aged 36-50 years (source: Health and Welfare Canada 1983).

households: households headed by a male adult, households headed by a female adult (in both cases with no spouse present), and three households headed by couples, where the wife works full time in the market, part time in the market, or full time in the home.

Per capita income and household size both affect food expenditure patterns. Consistent with theory, households with higher per capita income spend a lower share of their income on food, a higher share of food expenditures in restaurants, and more per calorie on home-pre-

pared food. Households headed by a single adult similarly allot a higher share of income to food and have a somewhat higher propensity to eat in restaurants and to buy more expensive home-prepared foods. In a univariate analysis, separating the effect of wife's employment status from the effects of income and household composition is difficult.

The nutrient availability results are of some potential concern. In eleven of the twelve cases, the households with the lowest apparent per capita intake are those where the wife works outside

Table 3. Regression Results for Food Share, Restaurant Share, and Cost per Calorie

Independent Variable	Food Share	FAFH Share	Cost per Calorie
Per capita income	-0.0137 (23.16)***	0.0107 (15.40)***	0.00241 (8.99)***
Per capita income squared	0.000112 (13.47)***	-0.0000723 (7.59)***	-0.0000205 (4.79)***
Number of adults 16-65	-0.0179 (5.26)***	-0.00758 (1.90)*	-0.000539 (0.41)
Number of adults 65 and over	-0.0266 (3.92)***	-0.0657 (8.26)***	-0.00219 (0.82)
Number of children	-0.0165 (5.09)***	-0.00633 (1.67)*	-0.00342 (2.67)***
Head has some post-secondary education	0.0193 (3.07)***	0.0351 (4.77)***	0.00796 (3.16)***
Head has university education	0.0299 (3.70)***	0.0243 (2.56)**	0.00858 (2.61)***
Female head, no husband	-0.00169 (1.87)*	0.0894 (8.42)***	0.00859 (2.54)**
Male head, no wife	0.0242 (2.33)**	0.109 (8.99)***	0.000381 (0.09)
Wife works full time in market	-0.00133 (0.16)	0.0397 (4.14)***	0.00679 (2.10)**
Wife works part time in market	-0.00264 (0.32)	0.0295 (3.04)***	0.00205 (0.63)
House owner	-0.00534 (0.90)	-0.0437 (6.30)***	-0.00705 (2.97)***
Quarter 2	-0.00750 (1.05)	0.00631 (0.75)	0.00904 (3.16)***
Quarter 3	0.00524 (0.73)	0.0272 (3.22)***	0.0173 (5.92)***
Quarter 4	0.00684 (0.95)	-0.000431 (0.05)***	0.000756 (0.26)
Maritimes	-0.0335 (4.06)***	-0.254 (2.62)***	-0.0112 (3.40)***
Quebec	0.00992 (1.20)	-0.0149 (1.54)	0.00219 (0.66)
West	-0.0149 (2.26)**	0.0216 (2.79)***	-0.00372 (1.42)
Adjusted R-squared	0.140	0.208	0.074
F-statistic	47.87***	76.53***	22.09***
Degrees of freedom	5,169	5,169	4,727
Intercept	0.371 (25.24)***	0.138 (8.01)****	0.145 (24.71)***

Notes: Figures in brackets below coefficients are *t*-statistics. Triple asterisk indicates significant at 1% level, one tail test; double asterisk indicates significant at 5% level, one tail test; single asterisk indicates significant at 10% level, one tail test. Coefficients are given to 3 significant figures, and *t*-statistics to 2 decimal places.

the home. This is somewhat surprising because these households have higher per capita incomes than households where the wife works full time in the home. Although households where the wife does not work in the market have the lowest per capita income of the five household types, in only one of the twelve cases does this household type have the lowest apparent per capita intake of a specific nutrient.

Tables 2, 3, and 4 contain the results of the multivariate analysis. Table 2 presents variable definitions, means, and standard deviations for the regressions, while tables 3 and 4 present the regression results: table 3 for the expenditure variables (share of food in income, share of

FAFH in food expenditures, and price per calorie of home-prepared food), and table 4 for the twelve nutrients.

The results for food expenditure, presented in table 3, generally confirm the findings from previous studies for other countries. Higher per capita income is associated with a decreased food share (Engels' law), but a shift occurs to more costly types of food as evidenced by an increase in FAFH share and a rise in the cost per calorie of home-produced food. The effect diminishes as income increases, as shown by the opposite sign of the squared terms. All the coefficients on income and income squared are significant. Household size has strong effects on food con-

Table 4. Regression Results for Apparent Nutrient Intake per Capita

Independent Variable	Dependent Variable					
	Calories	Protein	Fat	Carbohydrate	Folate	Calcium
Per capita income	26.984 (4.62)***	0.985 (4.45)***	1.217 (4.04)***	3.183 (4.21)***	2.406 (5.14)***	9.068 (3.93)***
Per capita income squared	-0.176 (1.89)**	-0.00660 (1.87)*	-0.00764 (1.59)	-0.0208 (1.73)*	-0.0133 (1.79)*	-0.0555 (1.51)
Number of adults	-176.656 (6.10)***	-4.971 (4.53)***	-8.392 (5.62)***	-21.248 (5.67)***	-15.214 (6.55)***	-70.897 (6.20)***
16-65						
Number of adults 65 and over	-94.985 (1.63)	-5.290 (2.39)**	-7.543 (2.51)**	-1.294 (0.17)	-10.242 (2.19)**	-80.495 (3.50)***
Number of children	-199.312 (7.16)***	-7.514 (7.12)***	-10.758 (7.50)***	-18.820 (5.23)***	-17.869 (8.01)***	-66.411 (6.04)***
Head has some post-secondary education	16.198 (0.30)	1.619 (0.78)	0.635 (0.23)	1.550 (0.22)	6.589 (1.50)	71.552 (3.31)***
Head has university degree	-20.373 (0.29)	0.134 (0.05)	-2.011 (0.55)	0.690 (0.08)	19.920 (3.47)***	107.161 (3.79)***
Female head, no husband	-169.825 (2.31)**	-6.492 (2.33)**	-11.945 (3.15)***	-8.755 (0.92)	-3.191 (0.54)	-123.981 (0.83)
Male head, no wife	141.693 (1.53)	7.119 (2.02)**	7.095 (1.49)	12.421 (1.04)	14.007 (1.89)*	101.611 (2.78)***
Wife works full time in market	-167.138 (2.38)**	-4.307 (1.62)	-7.948 (2.20)**	-20.228 (2.23)**	-8.946 (1.60)	-62.506 (2.25)**
Wife works part time in market	9.695 (0.14)	1.155 (0.43)	-1.702 (0.47)	4.873 (0.54)	1.330 (0.24)	-11.253 (0.40)
House owner	116.145 (2.25)**	5.460 (2.79)***	3.850 (1.45)	15.266 (2.29)***	6.809 (1.65)*	49.234 (2.42)**
Quarter 2	-40.260 (0.65)	-1.118 (0.47)	-2.189 (0.68)	-4.171 (0.52)	-6.709 (1.34)	-45.716 (1.86)*
Quarter 3	75.620 (1.19)	3.812 (1.59)	2.319 (0.71)	9.441 (1.15)	-4.165 (0.82)	-6.699 (0.27)
Quarter 4	114.339 (1.84)*	4.432 (1.88)*	4.042 (1.26)	14.982 (1.86)*	-6.849 (1.37)	-14.594 (0.59)
Maritimes	-28.835 (0.40)	-4.312 (1.58)	-6.289 (1.69)*	10.506 (1.13)	-8.651 (1.50)	-71.201 (2.50)**
Quebec	127.089 (1.77)*	4.477 (1.64)	4.231 (1.14)	18.274 (1.97)**	10.711 (1.86)*	-0.595 (0.02)
West	-98.654 (1.73)*	-6.159 (2.85)***	-4.952 (1.68)*	-7.760 (1.05)	-7.340 (1.60)	-66.284 (2.94)***
Adjusted R-squared	0.057	0.051	0.052	0.044	0.078	0.060
F-statistic	16.84***	15.14***	15.30***	12.98***	23.35***	17.82***
Degrees of freedom	4,727	4,727	4,727	4,727	4,727	4,727
Intercept	2594.687 (20.29)***	88.199 (18.19)***	121.873 (18.48)***	292.261 (17.67)***	212.666 (20.74)***	1098.643 (21.75)***

sumption: larger households allot a smaller share of income to food, a smaller share of the food budget to eating in restaurants, and buy cheaper food at home. There are three coefficients on household size variables in each of three equations: all nine are negative, seven of them significantly so.

House ownership is associated with a significantly lower restaurant share and a significantly lower cost per calorie of home-produced food. These results may reflect different preferences for home-based activities by owners and renters; nonowners may also face less adequate cooking facilities. More education of the household head is associated with a larger share of food in total

expenditures, a larger share of restaurant food in food expenditures, and a higher cost per calorie of home-prepared food. Two education dummy variables are in each of three equations, and all six of the coefficient are positive and significant. Education clearly has a large effect on tastes. Campbell and Horton, and Horton and Campbell examine these effects in more detail.

The variables of most interest in this study are those representing wife's employment. The reference group is households where the wife works full time in the home (once the norm for households, now representing just under a quarter of all households in the survey). Compared to this household type, all other households spend more

Table 4. Continued.

Vitamin A	Vitamin C	Iron	Thiamine	Riboflavin	Niacin
150.452 (4.13)***	1.759 (5.45)***	0.160 (4.22)***	0.0177 (4.07)***	0.0188 (3.77)***	0.473 (4.77)***
-1.045 (1.80)*	-0.0153 (2.99)***	-0.000638 (1.06)	-0.000160 (2.31)**	-0.0000949 (1.19)	-0.00296 (1.88)*
-829.788 (4.59)***	-8.297 (5.19)***	-1.119 (5.95)***	-0.102 (4.75)***	-0.143 (5.77)***	-2.410 (4.91)***
-257.101 (0.71)	-2.172 (0.99)	-0.455 (1.20)	-0.0542 (1.25)	-0.117 (2.34)**	-1.985 (2.00)**
-830.226 (4.78)***	-6.378 (4.15)***	-1.257 (6.97)***	-0.141 (6.80)***	-0.163 (6.86)***	-3.606 (7.65)***
366.344 (1.07)	7.424 (2.46)**	0.196 (0.55)	-0.0412 (1.01)	0.0310 (0.66)	0.520 (0.56)
782.188 (1.75)*	14.619 (3.70)***	0.387 (0.83)	-0.0852 (1.60)	0.0402 (0.66)	-0.445 (0.37)
153.589 (0.33)	4.273 (1.05)	-0.947 (1.99)**	-0.0932 (1.70)*	-0.0360 (0.58)	-2.739 (2.20)**
-503.903 (0.87)	1.346 (0.26)	0.943 (1.57)	0.0946 (1.37)	0.136 (1.71)*	2.569 (1.64)
-625.808 (1.43)	-1.650 (0.43)	-0.938 (2.06)**	-0.121 (2.33)**	-0.098 (1.63)	-1.987 (1.67)*
-44.709 (0.10)	-1.584 (0.41)	0.118 (0.26)	0.0198 (0.38)	0.0130 (0.22)	0.343 (0.29)
392.548 (1.22)	5.738 (2.01)**	0.835 (2.49)**	0.114 (2.98)***	0.137 (3.11)***	2.893 (3.31)***
-737.868 (1.90)*	-2.208 (0.64)	-0.217 (0.54)	-0.0260 (0.56)	-0.0688 (1.29)	-0.662 (0.63)
-100.693 (0.26)	0.388 (0.11)	0.290 (0.71)	0.0255 (0.54)	0.0548 (1.01)	1.069 (1.00)
-642.78 (1.65)*	-8.449 (2.46)***	0.459 (1.14)	0.0993 (2.14)**	0.0194 (0.36)	1.214 (1.15)
846.092 (1.88)*	-1.173 (0.30)	-0.416 (0.89)	0.0347 (0.65)	-0.0687 (1.11)	-2.052 (1.68)*
1703.046 (3.79)***	4.069 (1.02)	0.949 (2.03)**	0.086 (1.61)	0.0792 (1.29)	2.535 (2.08)**
-181.452 (0.51)	-6.864 (2.18)**	-0.410 (1.11)	-0.0560 (1.32)	-0.0785 (1.61)	-2.225 (2.30)**
0.036 10.80***	0.047 14.09***	0.058 17.09***	0.041 12.16***	0.050 14.81***	0.057 16.78***
4727 8376.302 (10.50)***	4727 91.784 (13.00)***	4727 15.563 (18.76)***	4727 1.685 (17.72)***	4727 2.144 (19.58)***	4727 40.055 (18.48)***

per calorie for home-prepared food, even controlling for per capita income; the effect is significant in two of the four cases. Other households also spend a significantly greater share of the food budget on FAFH. This result supports cost-of-time interpretations, i.e., households where the wife works full time in the home substitute the wife's time for purchased inputs in food preparation. Male-headed households with no wife present are likely to spend more on food as a proportion of their income. Despite changing sex roles in housework, food preparation still remains a primarily female responsibility. Men without wives are substantially less likely to cook for themselves and likely prepare more convenient, and usually more costly, items when they do cook. Men also have higher requirements for most nutrients than do women.¹⁰

Table 4 presents the results for apparent per capita intake of twelve nutrients. Although the number of significant coefficients is large, the explanatory power of the equations is not high. This is to be expected; food expenditure data (unlike nutrient intake data) are subject to changes in food stocks. Purchase of a large bag of flour or container of cooking oil in one week is likely to inflate substantially apparent nutrient intake. This stocking behavior cancels out over a large enough sample; however, it makes prediction of nutrient availability by an individual household a hazardous procedure.

Per capita income is associated with significantly increased apparent intake of all twelve nutrients but at a rate decreasing with income (the squared term is always negative, significantly so in 8 cases). Household size variables have a negative impact on nutrient availability in all thirty-six cases, significant and negative in thirty cases. The different sizes of the coefficients reflect different nutrient requirements by age. House owners, despite allotting a lower share of the budget to food and buying less expensive foods, have higher apparent per capita intake of all nutrients (10 out of 12 coefficients are significant). The education variables have interesting effects, suggestive of increased concern for health at higher levels of education. There is a positive effect on most nutrients, significant for vitamin C and calcium for both post-secondary and university education and significant for university education for vitamin A and folate.

However, university education has a negative effect on calories, fat, niacin and thiamine, and post-secondary education has a negative effect on thiamine (intake of the B vitamins is frequently correlated with calorie intake).

Finally, the results for the wife employment status variables are compared to the reference household (households headed by a couple, where the wife works full time in the home). Households headed by a female with no husband consume less of ten of the twelve nutrients, significantly less in six cases; women have lower requirements than men for most nutrients. Conversely, households headed by a male with no wife consume more of eleven of the twelve nutrients, significantly more in four cases. Households where the wife works full time in the market consume less of all twelve nutrients (significant in 8 cases) and households where the wife works part time consume less of four nutrients and more of eight (however, none of the effects are significant). Even holding constant household size, age composition, and income, there are very strong effects of wife's full time employment status on nutrient availability.

Because this conclusion is potentially contentious, some additional work was done on dietary adequacy (Campbell and Horton); those results are summarized here. Dietary adequacy norms for Canada exist for ten of the twelve nutrients (the exceptions are fat and carbohydrates). These recommended nutrient intakes (RNI's: Health and Welfare Canada 1983) are similar to the U.S. RDA's. Nutrient requirements were calculated for each household separately, using the household age composition information. As discussed earlier, the family food expenditure survey data has only fairly broad age categories and no sex composition data. Therefore, nutrient requirements were constructed for these broad age categories using the 1986 census population weights. Households were then categorized as having adequate or inadequate diets for each of the ten nutrients.

The results of this analysis found that a non-trivial fraction of diets were classified as inadequate, ranging from a low of 6.2% (vitamin A) to a high of 52.4% (folate). Multivariate analysis was undertaken to estimate the probability that the household had an inadequate diet for a particular nutrient.¹¹ The same independent variables were used as before, and probit meth-

¹⁰ The tobit results for the FAFH share are fairly similar to the OLS ones, although some of the coefficients on the region and quarter dummies change sign. Note that the standard errors in the appendix table are not exactly correct; they have not been corrected for the fact that instrumental variables were used in conjunction with the tobit estimates.

¹¹ It has been suggested that adequacy measures for single nutrients are less meaningful than adequacy measures for groups of nutrients. Pete-kin, Kerr and Hama suggest an adequate diet should contain 80% of the RDA's for 11 nutrients, and Huffman suggests it should contain 100% of the RDA's for 13 nutrients.

ods were used to estimate the probability that a household had an adequate diet. A negative effect of wife's full-time employment was found for all ten nutrients, significant in eight cases. Thus, the adverse effects of wife's employment status cannot be dismissed by arguing that the effect exists only for households who are already exceeding their nutrient requirements.

It is possible that the RNI's are set rather high and that households with working wives are restricting their intake for health reasons. Although limiting intake of calories, fats, and carbohydrates might seem desirable, this is less clear for calcium or iron. These households might compensate by consuming more vitamin and mineral supplements, and it would be desirable to explore this possibility in future food expenditure surveys.

Finally, although the effect of the wife's higher cost of time might reduce nutrient intake, this is offset by the higher per capita income of households with wives employed in the market. As shown in table 1, however, the effect of higher income is not large enough to offset the other adverse effects, and households with wives working outside the home end up with lower nutrient availability.

Conclusions

This study has analyzed the behavior of a set of summary variables describing food expenditures and nutrient availability (3 expenditure variables and 12 nutrient variables) for a large national sample of urban households in Canada. As in previous studies, income, household size, and house ownership have important effects, and higher levels of education are correlated with increased attention to health consideration in food selection. Wife's employment status, via its effects on cost of time, leads to a higher share of restaurant food consumption and a higher cost per calorie of home-prepared food (similar to findings of McCracken and Brandt, and Kinsey, for the U.S.).

There are negative effects on nutrient availability for households where the wife works full time in the market. The higher per capita income available to these households does not offset the cost-of-time effect, and there is no evidence that higher wife's income leads to increased expenditures on "basic needs." Out of the five household types, households with wives employed in the market (either full and part time) have the lowest apparent mean intakes of ten of the twelve nutrients. Work elsewhere by the au-

thors showed that wife's employment status had a significant effect on increasing the likelihood of dietary inadequacy for several nutrients. This contrasts with findings for the United States (Adrian and Daniel), which were that wife's market work was associated with increased intake for six out of eight nutrients studied (significant for 2 nutrients). Methodological differences between the two studies may be an issue because Adrian and Daniel use household rather than per capita levels for variables, and they also have wife's employment status as exogenous.

Because there are some limitations of using expenditure data to examine nutrient availability, further analysis with more ideal data would be potentially useful, in particular including information on the foods eaten away from home. The study also indicates that U.S./Canadian comparisons would be a fruitful topic for further study, particularly given the past underutilization of the Canadian data.

Finally, regarding the policy implications: the authors are very far from suggesting that wives ought to be persuaded not to work outside the home. The results, however, do confirm the stresses borne by women and their households when women work outside the home and end up with a "double day" of work. Policies designed to help households deal with these stresses might be useful. This would be especially important for low income households where both adults work. In the food area, these policies might include augmenting existing efforts to ensure the optimal nutrient content of restaurant meals (including fast food meals, and meals eaten away from home by children) and nutrition education efforts for nutrient-dense, convenient, home food preparation techniques. In addition to policy changes, major changes in societal values affecting the division of labor within the household (such as involving men more in food purchasing and preparation) may also ease some of the stresses experienced by households where wives work outside the home, reflected in these results.

[Received May 1989; final revision received October 1990.]

References

- Adrian, J., and R. Daniel. "Impact of Socioeconomic Factors on Consumption of Selected Food Nutrients in the United States." *Amer. J. Agr. Econ.* 58(1976):31-38.
- Barewal, S. "Canadian Spending of Foodservice Dollars." *Food Mkt. Commentary* 9(1987):32-41.

- Campbell, C., and S. Horton. "Apparent Nutrient Intakes of Canadians: Continuing Nutritional Challenges for Public Health Professionals." *Can. J. Public Health*, in press.
- Capps, O., Jr., J. R. Tedford, and J. Havlicek, Jr. "Household Demand for Convenience and Nonconvenience Foods." *Amer. J. Agr. Econ.* 67(1985):862-69.
- Deaton, A., J. Ruiz-Castillo, and D. Thomas. "The Influence of Household Composition on Household Expenditure Patterns: Theory and Spanish Evidence." *J. Polit. Econ.* 97(1989):170-200.
- Dwyer, D. H. "Women and Income in the Third World: Implications for Policy." New York: Population Council International Program Work. Pap. No. 18, 1983.
- Health and Welfare Canada. *Recommended Nutrient Intake for Canadians*. Ottawa: Minister of Supply and Services Canada, 1983.
- Horton, S., and C. Campbell. "Do the Poor Pay More for Food?" *Food Mkt. Commentary* 11(1989):33-39.
- Hunt, L. "Socioeconomic Profile and Food Expenditures of the Single-Parent Family in Canada, 1974-82." *Food Mkt. Commentary* 7(1985):55-63.
- Karamchandani, and L. Robbins. "Changes in Food Expenditure Patterns 1969-76." *Food Mkt. Commentary* 1(1979):34-37.
- Kennedy, E., et al. "Distribution of Nutrient Intake Across Meals in the United States Population." *Ecology of Food and Nutrition* 11(1982):217-24.
- Kinsey, J. "Working Wives and the Marginal Propensity to Consume Food Away from Home." *Amer. J. Agr. Econ.* 65(1983):10-19.
- Lee, L. F., and M. M. Pitt. "Microeconomic Demand Systems with Leading Non-negativity Constraints: The Dual Approach." *Econometrica* 54(1986):1237-42.
- Mathieson, A., and L. Robichon-Hunt. "Food Expenditure Patterns and Apparent Nutrient Intakes with Particular Reference to Low Income Families in Canada 1974-78." Mimeographed. Ottawa: Agriculture Canada Food Research Institute, 1985.
- McCracken, V. A., and J. A. Brandt. "Household Consumption of Food-Away-from-Home: Total Expenditure by Type of Food Facility." *Amer. J. Agr. Econ.* 69(1987):274-84.
- Morgan, K. J. "Socioeconomic Factors Affecting Dietary Status: An Appraisal." *Amer. J. Agr. Econ.* 68(1986):1240-46.
- Nelson, M., P. A. Dyson, and A. A. Paul. "Family Food Purchases and Home Food Consumption: Comparison of Nutrient Contents." *British J. Nutrition* 54(1985):373-87.
- Nevraumont, U. "Where Canadians Buy Their Food and When." *Food Mkt. Commentary* 9(1988):48-53.
- Robbins, L. "The Food Expenditure Patterns of the Elderly in Canada, 1974-82." *Food Mkt. Commentary* 8(1986):29-43.
- Senauer, B., D. Sahn, and H. Alderman. "The Effect of Value of time on Food Consumption Patterns in Developing Countries: Evidence from Sri Lanka." *Amer. J. Agr. Econ.* 68(1986):920-27.
- Sexauer, B. "The Effect of Demographic Shifts and Changes in the Income Distribution on Food-Away-from-Home Expenditure." *Amer. J. Agr. Econ.* 61(1979):1046-57.
- Statistics Canada. *Family Food Expenditure in Canada: Selected Cities, 1984*. Ottawa: Minister of Supply and Services Canada, 1986.
- Tucker, K., and D. Sanjur. "Maternal Employment and Child Nutrition in Panama." *Soc. Sci. and Medicine* 26(1988):605-12.
- Wales, T. J., and A. D. Woodland. "Estimation of Consumer Demand Systems with Binding Non-negativity Constraints." *J. Econometrics* 21(1983):263-85.

Appendix

Tobit Estimate Results for FAFH Share

Independent Variable	FAFH Share
	0.0119
Per capita income	(24.74)***
	-0.000907
Per capita income squared	(18.29)***
	-0.00500
Number of adults 16-65	(1.47)
Number of adults 65 and over	-0.0278
	(4.77)***
	-0.00878
Number of children	(2.21)**
Head has some post-secondary education	0.0353
	(5.31)***
Head has university education	0.0220
	(2.47)**
	0.0887
Female head, no husband	(11.13)***
	0.109
Male head, no wife	(21.09)***
Wife works full time in market	-0.159
	(2.48)***
Wife works part time in market	0.600
	(7.49)***
	-0.375
House owner	(6.04)***
	0.0122
Quarter 2	(1.60)
	0.0332
Quarter 3	(4.43)***
	0.00695
Quarter 4	(0.90)
	-0.0397
Maritimes	(4.85)***
	-0.0210
Quebec	(2.83)***
	-0.0169
West	(2.02)**
Log likelihood	-1353.1
Sample size	5,188

Long-Run versus Short-Run Planning Horizons and the Rangeland Stocking Rate Decision

L. Allen Torell, Kenneth S. Lyon, and E. Bruce Godfrey

Reduced future forage production, diminished range condition, and reduced animal performance have been major factors considered when setting rangeland stocking rates. The relative economic importance of diminished current period animal performance versus intertemporal forage production impacts was investigated using a dynamic optimal control model. The model is applied to yearling stocker production in eastern Colorado. Results indicate that intertemporal grazing impacts to forage production are not that important; reduced weight gain during the current period drives the economic stocking rate decision. Further, ranchers have no economic incentive as profit maximizers to continually overgraze the range.

Key words: livestock grazing, optimal control, optimal rangeland use, range economics, stocking rates.

Overgrazing can reduce future forage flows, as evidenced by the history of livestock grazing on much of the public domain (Stevens and Godfrey). Because of this, setting rangeland stocking rates is considered one of the most important grazing management decisions from the standpoint of vegetation, livestock, wildlife, and economic returns (Holechek, Pieper, and Herbel). In addition to the past common property problem (Hardin), the short-term planning horizon of livestock producers, even when property rights are assigned, is often cited as the reason for rangeland degradation (Workman, p. 52). Some argue that the overgrazing problem would be solved if ranchers would only consider the impact of current stocking decisions on future range condition and production.

L. Allen Torell is an associate professor of agricultural economics, New Mexico State University, Las Cruces; Kenneth S. Lyon is a professor, and E. Bruce Godfrey is an associate professor, Department of Economics, Utah State University, Logan.

This is Journal Article No. 1510 of the Agricultural Experiment Station, New Mexico State University.

This research derives from and extends Torell's doctoral thesis, which was supported by the Utah State University Agricultural Experiment Station. The empirical model application was conducted by William W. Riggs as a master's thesis at New Mexico State University.

The authors appreciate helpful reviews of earlier drafts by John P. Workman, Utah State University; Richard H. Hart, U.S. Department of Agriculture, ARS High Plains Grasslands Research Station; Soumen Ghosh, New Mexico State University; and the anonymous *Journal* referees.

Similar to the apparent short-term planning horizon of some ranchers, past economic stocking rate models have also been developed from a static, short-term perspective (see for example, Hart et al. 1988a,b; Hildreth and Riewe; Torell and Hart; Workman). These models are driven by falling animal performance (i.e., weight gain, calf crop, conception rates) as stocking rates increase, and the models ignore the impact of grazing on future range condition and production.

Stocking rates determined to be economically optimal over the single period may not be optimal at all when dynamic forage production impacts are considered. The relative importance of stocking rates in determining current period animal performance and the dependence of resource flows on prior use rates will determine the need for sophisticated dynamic models when making stocking rate decisions.

In this paper, a dynamic economic model that incorporates both current and future period impacts of the stocking rate decision is developed. The model can be used to investigate the optimal frequency of rangeland investments, and the investment of judicious grazing practices to extend and enhance rangeland productivity is directly considered. Similar dynamic stocking rate models have been developed by Pope and McBryde, Karp and Pope, and Rodriguez and Taylor; but our application is unique in that op-

timal stocking rates determined by the dynamic model are compared to economic stocking rate recommendations based on traditional single-period (myopic) stocking rate models. Through this comparison, the relative importance of grazing impacts to current and future period livestock production are examined and the need for more sophisticated dynamic models for stocking rate decisions is evaluated.

The economic model is developed and applied to stocker cattle operations because data were available and because this allows certain simplifications to be made. However, stocking rates for stocker cattle operations are substantially more adjustable through time in response to forage availability and cost/price variations than are cow/calf enterprises. The incentive for ranchers' to retain brood stock through drought periods and to incur short-term losses, so as to avoid the expense of buying additional brood stock at a later date, is an important consideration that may increase the motivation for cow/calf producers to overgraze.

The Traditional Single-Period Stocking Rate Model

The traditional value of marginal product model of optimal input use has been applied to the stocking rate problem by numerous authors (Hart et al. 1988a,b; Hildreth and Riewe; Riewe; Torell and Hart; Workman). Although the variable input differs, these economic evaluations start with a basic definition of input/output relationships. Most recently, stocking rate studies have standardized the grazing input to grazing pressure (GP is animal days of grazing per unit of forage) or forage allowance (animals/unit of forage) (Hart et al. 1988a, Vallentine). This standardization adjusts for grazing intensity differences as forage production varies between years because of weather and other environmental factors.

We use the basic definitions of Hart et al. (1988a) and define stocking rate (SR) as the number of stockers grazing per hectare over a grazing period of length v , and SD as the number of stocker days grazing per hectare (i.e., $SD = v \cdot SR$). GP is current period grazing pressure measured in stocker days of grazing per unit of standing herbage (H) produced (metric ton = 1,000 kilograms—kg);

$$(1) \quad GP = SD/H = v \cdot SR/H.$$

In this single-period model, herbage production

and the length of the grazing period are exogenously defined (or estimated) at the time the stocking rate decision is made. This assumed knowledge of peak standing forage production for the year would be limiting in some cases, especially when warm-season grasses predominate the range site and the majority of grass growth is not realized until well after the range is stocked.¹ At other locations, where cool-season (early growing) grasses predominate, a satisfactory estimate of expected peak standing forage crops could be made. For example, Hart found for cool season grasses growing on the High Plains of Wyoming, annual forage production had nearly peaked by the time pastures were stocked in May, with 94% of the variation of annual forage production explained by the amount of March–May precipitation.

The relationship between gain per animal per day (average daily gain, ADG) and SR is defined to be a concave function given by

$$(2) \quad ADG = f(GP(SR)),$$

with $df/dSR < 0$ and $d^2f/dSR^2 \leq 0$ over the economically relevant range of production.

The total kilograms of beef produced per hectare, as given by the production function, is defined by multiplying the number of steers grazing per hectare times their average sale weight (W_s);

$$(3) \quad b(GP(SR)) = SR \cdot W_s$$

equals kilograms of beef sold per hectare, where

$$(4) \quad W_s = [W_p + v \cdot f(GP(SR))]$$

equals stocker sale weight (kg), and W_s is determined by the stocking rate decision. For a given level of herbage production, increased stocking rates will decrease ADG and average sale weights. Average sale weight will also depend on purchase weight (W_p) and the length of the grazing period (v).

Heavier feeder cattle generally sell for less (Schroeder et al.). While sale price is determined by market forces outside the rancher's control, the rancher determines which market price to accept by the size and quality of cattle produced. This depends on the stocking rate decision such that, in addition to market forces, P_s

¹ Rodríguez and Taylor show that it would be economically advantageous to stock heavily but maintain a flexible marketing strategy and sell early if realized forage conditions were less than anticipated. This would be equivalent to adjusting v so as to maintain the optimal GP .

is a function of SR and other livestock characteristics;

$$(5) \quad P_s = P_s(W_s(GP(SR)), X),$$

where X is a vector of exogenous variables that identify relevant characteristics of the stocker cattle at time of sale (e.g., breed, frame size, health, fill, sex, muscling) and price expectations.

The rancher selling in the competitive market must choose the stocking rate that maximizes profit where, like the model of Hildreth and Riewe, profit per hectare is defined to be

$$(6) \quad \pi(GP(SR)) = P_s(W_s(GP(SR)), X) \cdot b(GP(SR)) - (P_p W_p + r) SR - a,$$

where P_p is steer purchase price (\$/kg); $P_p W_p$ is per head animal purchase cost, r is per head input factor costs (e.g., supplemental feed, vaccines, labor), which vary with the number of stockers grazing the pasture, and a is total fixed production costs, which do not vary with use rate of the pasture.

Differentiating equation (6) with respect to SR gives the standard result for optimal input use; at the optimum, the value of the marginal product (VMP) from adding another stocker to the pasture will equal the marginal factor cost (MFC) of putting it there. Considered in the first-order condition are the marginal value and marginal cost of adding the last stocker to the pasture, the negative impact to ADG of all animals in the pasture, and adjustments to expected selling price as sale weights change with stocking rate.

The solution to the myopic stocking rate model may not be optimal when dynamic impacts of current stocking rates on future forage production are considered. With a dynamic model formulation, the significant change is that stocking rate decisions during the current period directly impact forage production, potential beef production, and revenue in all future years.

A Dynamic Stocking Rate Model

In a dynamic setting, a rancher must choose the rangeland use rates that maximize the present discounted value of economic returns calculated over a multiperiod planning horizon. If rangeland improvement practices will rejuvenate the productivity of a range site depleted by the past time path of stocking rate decisions or from naturally occurring brush invasion not related to grazing, a chain of renewal cycles (range im-

provements) exists. In this case, renewal cycle S , defined as the length of time between two pasture improvements, must be chosen. If no feasible rangeland improvement practice can be implemented, the current stand of grass must be managed from now to infinity.

To model the interaction between use rates and resource flows, we follow the same procedure outlined by Stevens and Godfrey. An index of rangeland productivity at time t is defined to be $I(t)$ ($0 \leq I(t) \leq 1$).² This index is a measure of average rangeland productivity obtained with the grazing policy being evaluated, relative to average production sustainable through time under light or no grazing (\bar{H}).

The average rangeland productivity index changes through time, depending on the grazing history of the pasture. If relatively heavy stocking occurred last year, the herbage production index this year may be reduced. Movement of the index is defined to be a function of the previous period's GP , the previous period's index value, and additional productivity declines not related to grazing use rates (e.g., naturally occurring brush invasion and diminished site productivity).

Average herbage production at a point in time is given by the equation

$$(7) \quad H(t) = I(t)\bar{H},$$

with past impacts of grazing captured by $I(t)$.

By differentiating equation (7), the equation of motion for the state variable $H(t)$ is given by a differential equation of the form

$$(8) \quad dH/dt = dI/dt \cdot \bar{H} = \Phi(GP(SR(t)), H(t), t).$$

With the initial condition that at the start of the planning period, $I(t_0) = 1$ or $H(t_0) = \bar{H}$, equation (8) defines the time path of annual average herbage production, depending on the solution set to $SR(t)$.

Average herbage production would be expected to decline through time as brush invades the pasture. Further, heavier stocking rates would be expected to accelerate the rate of brush invasion and the decline in rangeland productivity; this grazing interaction is captured in the dynamic model. $H(t)$ is an endogenously determined state variable, and the time path of herbage pro-

² Normally, the maximum index value might be set at one, implying light or no grazing yields maximum herbage production. However, it is possible that with alternative grazing systems a heavier stocking rate would yield higher herbage production through time and thus give an index greater than one.

duction depends on the stocking rate decisions made over the planning period.

Random fluctuations in weather conditions, especially rainfall, will cause annual deviations from the herbage production defined by equation (7), but these deviations will be offsetting on average. The computed average herbage production from equation (7) is the expected herbage production at that point if random variation in herbage production is normally distributed about some mean production level.³ This mean production level will change depending on stocking rate decisions, and this changing average production is tracked in the dynamic model.

The annual profit function of equation (6) is now altered to include herbage production:

$$(9) \quad \pi(GP(SR(t), H(t))) \\ = P_s(W_s(GP(SR(t), H(t))), X) \\ \cdot b(GP(SR(t), H(t))) \\ - (P_p W_p + r)SR(t) - a.$$

The profit-maximizing rancher would maximize the discounted net present value (NPV) from grazing over all future years. The wealth from a single grass stand rotation is given by

$$(10) \quad C(1) = \int_{t_0}^S \pi(GP(SR(t), H(t)))e^{-\rho t} dt - K(0),$$

where ρ is the discount rate used to discount future returns, t is the year under consideration, and $K(0)$ is the cost of stand rejuvenation implemented at time 0 (e.g., brush control, reseed, pasture reestablishment).

The integral is defined to start at time t_0 ($0 \leq t_0 < S$), which reflects the specified grazing deferment policy associated with grass stand rejuvenation. Assuming all rotations are alike, following Perrin's replacement model, the wealth from all future rotations is given by

$$(11) \quad C(\infty) = A(S)C(1),$$

where $A(S) = 1/(1 - e^{-\rho S})$ = infinite series factor, the present value of a perpetual annuity received every S years.

The objective functional given by $C(\infty)$ is

maximized subject to the initial stock of herbage available for grazing during period t_0 ($H(t_0) = \bar{H}$), and the law of motion for the system as given by equation (8). If rejuvenation of the pasture is possible, the optimal rotation period (S) must also be chosen. If rejuvenation is not possible, such that S is set to ∞ , maximizing $C(1)$ or $C(\infty)$ is equivalent.

Using theorem 1 of Long and Vausden, the Hamiltonian for the maximization problem can be written as

$$(12) \quad H = A(S)[\pi(GP(SR(t), H(t)))e^{-\rho t} - K] \\ + \omega \cdot \Phi(GP(SR(t), H(t))).$$

Given equation (8), this is an optimal control problem, where S is a control parameter, $SR(t)$ is the control variable (function), $H(t)$ is the state variable, and ω is the costate variable defining the shadow value of herbage at each point in time.

In both the static and dynamic model, grazing pressure directly impacts current period beef production through grazing pressure impacts to ADG. However, in the dynamic model, the effect of stocking intensity in previous years is captured through its impact on herbage production during the current period. Profits at a point in time are a function of $SR(t)$ and the amount of herbage at that point. Average herbage production evolves from starting point \bar{H} , depending on how heavily the pasture is grazed. Two implicit assumptions are (a) pasture rejuvenation (e.g., killing the brush, reseeding) followed by the specified deferment policy will result in the return of grass production to the beginning level given by \bar{H} , and (b) each treatment cycle is identical.

Similar to the replacement criterion developed by Perrin and by Chavas, Kliebenstein, and Crenshaw, the optimal rotation period is extended to the point where the rent during the S th period is equal to the interest that could be earned on the wealth obtained by optimally utilizing the forage over the S -year rotation period.

Model Application

The maximization problem so far has been represented with continuous-time variables to allow simpler algebraic analysis. We illustrate a discrete time optimal control formulation of the dynamic problem below and use this empirical model to determine the time path of rangeland productivity that would result under profit-max-

³ Using a similar dynamic stocking rate model and simulated data from the SPUR rangeland simulation model (Wight), Torelli investigated how weather variations affected the dynamic stocking rate decision. Optimal GP and the time path of herbage production were found to be similar over a wide range of alternative weather patterns.

imizing behavior by ranchers. For the dynamic model, we solve for the optimal time path of $SR(t)$, for a number of rotation lengths, then select the grass stand replacement age that is optimal, as indicated by the maximum objective functional value $[C(\infty)]$. Solution to the empirical model was obtained using the MINOS nonlinear programming algorithm (Murtaugh and Saunders).

To evaluate differences in optimal use rates and in optimal rangeland productivity through time, the traditional single-period (myopic) model was iteratively solved for similar price/cost situations. This was done by determining the myopic optimal stocking rate for year 1 and then using the estimated equation of motion to define average herbage production that would result in year 2 with this SR . The optimal stocking rate for year 2 was then calculated and the procedure was repeated over the planning horizon. A similar procedure, but with dynamic interactions considered in the optimization, was also used for the dynamic model. This comparison of the myopic and dynamic solutions indicates the importance of developing more complex dynamic models and provides an evaluation of diverging trends in rangeland productivity when the future is not considered in stocking rate decisions.

The Study Site

The empirical analysis uses data from a long-term grazing study conducted in eastern Colorado (see Sims, Dahl, and Denham for a complete description of the grazing trial methods and results). This Colorado example is one of the few grazing studies with adequate design and length to quantify the long-term impacts of grazing to livestock and forage production needed for economic analysis.

The Colorado grazing study was conducted over fifteen years, from 1955 through 1969, with both livestock and forage response to grazing treatments measured during selected years. The study site was in the central Great Plains, in the northeastern Colorado sandhills. Average annual precipitation over the study period was about 38 centimeters (15 inches). The distribution of rainfall was normally distributed, with an equal number of years below and above the reported average. About 73% of the precipitation occurred during the active growing season between 1 April and 31 August.

The vegetation was heterogenous with major forage species of blue grama (*Bouteloua gra-*

cilis), prairie sandreed (*Calamovilfa longifolia*), and needle-and-throat (*Stipa comata*). The primary invading brush species was sand sagebrush (*Artemisia filifolia*).

Three continuous season-long grazing treatments were implemented during 1955, with 4 hectares allowed per grazing steer for the defined light stocking rate, 2 hectares for the moderate stocking rate, and 1.3 hectares for the heavy rate. The grazing season was 150 days long, from about 1 May to 30 September. Livestock response data were recorded from 1955 through 1965, and herbage production data were recorded from 1957 through 1968. Average herbage production measured over the twelve years was 1,728, 1,485, and 1,311 kilograms per hectare under light, moderate, and heavy grazing treatments, respectively.

Table 1 shows the herbage production and percent forage utilization reported for each of the three stocking rates. Also shown is the grazing pressure calculated for each year of the grazing trial, based on other data reported by Sims, Dahl, and Denham. The data reported in table 1, along with other livestock response data reported in the Colorado grazing study were used to estimate the equations of the optimization model.

Rangeland productivity was negatively impacted through time by increased grazing pressure. In fact, the Colorado grazing study found, for this range site, that any rangeland productivity deterioration was related directly to grazing pressure.⁴ Rangeland productivity increased slightly under the light grazing treatment, especially during the first seven years of the study. It decreased under the heavy use rate and remained relatively constant at the moderate rate. By 1965, when the light grazing treatment was discontinued, herbage production under moderate grazing averaged about 94% that of the light treatment, compared to 82% under the heavy treatment.

In addition to reduced herbage production, increased grazing intensity had adverse, but not major, impacts to the range resource; it de-

⁴ This result is in contrast to the range replacement model of Burt (1971), where grazing was assumed to have no significant impact on the rate of brush encroachment and productivity decline. In response to criticism on this assumption, Burt (1972) stated that there is little empirical evidence to support or refute the importance of grazing to the retreatment problem. We would agree with Burt's observation for the piñon-juniper problem he addressed; essentially, no empirical work has been done on this problem since Burt's paper was written. As a result, grazing impacts to future resource flows are largely unquantified and undoubtedly range site and species specific.

Table 1. Herbage Production (*H*), Grazing Pressure (*GP*) and Forage Utilization Under Three Stocking Rates, 1957–68

Year	Stocking Rate										
	Light			Moderate				Heavy			
	GP	H(t)	Utility	GP	H(t)	I(t) ^a	Utility	GP	H(t)	I(t) ^b	Utility
		(kg/ha)	(%)		(kg/ha)		(%)		(kg/ha)		(%)
1957	26	1,339	20	48	1,432	1.07	41	72	1,497	1.12	62
1958	23	1,752	26	46	1,762	1.01	41	74	1,691	0.96	63
1959	31	1,423	43	62	1,448	1.02	56	105	1,312	0.92	79
1960	24	1,738	29	55	1,521	0.88	34	99	1,316	0.76	57
1961	18	1,999	29	40	1,794	0.90	34	64	1,618	0.81	57
1962	18	2,005	35	45	1,546	0.77	32	67	1,604	0.80	62
1963	18	2,092	49	41	1,523	0.73	41	60	1,499	0.72	50
1964	23	1,662	19	54	1,399	0.84	50	98	1,206	0.73	70
1965	24	1,547	20	50	1,461	0.94	40	90	1,272	0.82	58
1966					1,477		49		1,264		63
1967					1,209				920		
1968					1,250				532		
1957–65											
Average	23	1,728	30	49	1,485	0.91	44	81	1,311	0.76	64

Source: Sims, Dahl, and Denham.

^a Herbage production under moderate grazing divided by herbage production under light grazing.

^b Herbage production under heavy grazing divided by herbage production under light grazing.

creased the amount of desirable cool season grasses, diminished range condition, and increased the density and overstory of sand sagebrush.

As in many other grazing studies, average daily gain and gain per head were lowest under heavy stocking rates. Holechek, Pieper, and Herbel attribute this decline in livestock performance to reduced forage intake and diet quality. Decreased forage availability reduces animal selectivity and forces grazing animals to select diets lower in quality. It also forces animals to spend more energy on foraging activities that could otherwise go into production.

Beef production per hectare is determined by the weight of the animals and the number of animals in the pasture. In the Colorado grazing study, beef produced per unit land area was highest under the heaviest stocking rate for every year of the study.

Model Specification

Using data from Sims, Dahl, and Denham, various model equations and relationships, including the production function and equation of motion, were estimated using ordinary least squares regression. The beef price model of Schroeder et al. was used to determine selling price as a function of the endogenously determined sale weight. In addition, the linear relationship between *GP* and forage utilization was estimated

so as to relate our model results to the standard stocking rate prescriptions made by range scientists using percent forage utilization as an indicator of proper use.⁵ Estimated regression equations are shown in table 2.

Beef production function. Several different functional forms of the average daily gain function have been specified in past grazing research, but a linear equation has been the most common specification and adequately defines *ADG* response to grazing pressures (Bransby et al.; Hart et al. 1988a,b; Hildreth and Riewe). With a linear *ADG* function, equations (3) and (4) imply a quadratic production function.

Equation of motion. The equation of motion was estimated as a linear first-order difference equation. We tried a number of different functional forms of the equation, but a linear equation seemed to fit best. As shown in table 1, there was no trend of herbage production that was not associated with the level of grazing use. Under light grazing, herbage production, range condition and the amount of brush in the pasture was nearly constant through time.

⁵ Current stocking rate recommendations, for example, are to maintain rangeland productivity by harvesting no more than 30% to 60% of available forage, depending on range type (Holechek, Pieper, and Herbel, p. 190).

Table 2. Selected Equations of the Stocking Rate Models

Average daily gain (kg/head/day) ^a		
$ADG = f(GP_t) = 0.82 - 0.0029 \cdot GP_t$		
Sale weight (kg/head) ^{a,b}		
$W_s = [W_p + v \cdot ADG] = 338 - 0.436 \cdot GP_t$		
Livestock gain per hectare (kg/ha) ^b		
Gain = $SR_t \cdot v \cdot ADG = SR_t(123 - 0.436 \cdot GP_t)$	(6.2673) ^c (0.0845)	$R^2 = 0.99$
Beef production per hectare ^b		
$b(GP(SR)) = SR_t \cdot W_s = SR_t(338 - 0.436 \cdot GP_t)$		
Equation of motion		
$I_t = 0.4343 + 0.5824 \cdot I_{t-1} - 0.00136 \cdot GP_{t-1}$	(0.1222) (0.1169) (0.0005)	$R^2 = 0.72$
$H_t = I_t \cdot \bar{H}$		
Percent forage utilization		
$U_t = 17.8008 + 0.5220 \cdot GP_t$	(3.582) (0.0628)	$R^2 = 0.73$
Beef prices ^d		
$P_p = 1.5064 + 0.465 \cdot Fut_{spring} - 0.00348 \cdot W_p + 0.000002 \cdot W_p^2$		
$P_s = 0.5626 + 0.314 \cdot Fut_{fall} + 0.00238 \cdot W_s - 0.000004 \cdot W_s^2$		

^a The equation for livestock gain per hectare was estimated directly from data presented by Sims, Dahl, and Denham. The *ADG*, sale weight, and beef production per hectare functions were algebraically estimated from this equation.

^b Assumptions: $W_p = 215$ kg, and $v = 150$ days.

^c The standard error of the estimate is presented in parentheses.

^d Reduced equations from Schroeder et al. with the livestock characteristics defined.

The herbage production index was estimated as

$$(13) \quad I_t = \alpha_0 + \alpha_1 I_{t-1} + \alpha_2 GP_{t-1}.$$

Then, I_t was multiplied by $\bar{H} = 1,728$ kilograms per hectare, the average herbage production under light grazing (table 1), to give

$$(14) \quad H_t = (\alpha_0 \bar{H} + \alpha_1 H_{t-1} + \alpha_2 \bar{H} \cdot GP_{t-1}).$$

In solving the optimization problem, the NPV of annual profits was maximized subject to constraints defining the dynamics of the system using equation (14). The constraint set was defined to be

$$(15) \quad \begin{matrix} A & X & = & b \end{matrix}$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ -\alpha_1 & 1 & 0 & 0 & \dots & 0 & -\alpha_2 \bar{H} & 0 & \dots & 0 \\ 0 & -\alpha_1 & 1 & 0 & \dots & 0 & 0 & -\alpha_2 \bar{H} & \dots & 0 \\ \vdots & & & & & \vdots & & & & \vdots \\ 0 & 0 & 0 & 0 & \dots & -\alpha_1 & 0 & 0 & \dots & -\alpha_2 \bar{H} \end{bmatrix} \begin{bmatrix} H_{t_0+1} \\ H_{t_0+2} \\ H_{t_0+3} \\ \vdots \\ H_{t_0+S} \\ GP_{t_0+1} \\ GP_{t_0+2} \\ \vdots \\ GP_{t_0+S} \end{bmatrix} = \begin{bmatrix} \bar{H} \\ \alpha_0 \bar{H} \\ \alpha_0 \bar{H} \\ \vdots \\ \alpha_0 \bar{H} \end{bmatrix}$$

⁶ If range site productivity declined over time from additional naturally occurring brush invasion and maturation that was not related to grazing, equation (13) could be specified with an additional time variable such as $I_t = \alpha_0 + \alpha_1 I_{t-1} + \alpha_2 GP_{t-1} + \alpha_3 t$, ($\alpha_3 < 0$).

Beef price function. Depending on livestock characteristics at the time of purchase and sale, the beef price model of Schroeder et al. can be used to estimate purchase and sale prices for feeder cattle. As shown by these authors, heavier feeder steers generally sell for less per kilogram. Thus, an increased stocking rate would be expected to increase sale price.

The Schroeder et al. model uses observed futures market prices to form price expectations. In this application, the futures prices in the fall (Fut_{fall} , on 1 October) and spring (Fut_{spring} , on 1

May) over the period 1979 to 1988 were used. This ten-year cycle of futures prices was repeated to estimate expected purchase and sale

prices of steers over a specified forty-year planning horizon.⁷

We assumed that uniform lots of healthy, mixed breed, medium sized cattle would be bought and sold. These assumptions result in reduced parameter price equations, with only the exogenously defined futures price and endogenous livestock sale weight as explanatory variables (table 2). The purchase price of stockers was exogenously defined with specification of the spring futures price and an assumed 215-kilogram purchase weight.

Table 3 shows the ten years of estimated beef prices used in the analysis, assuming an average 310-kilogram sale weight of steers.⁸ The smaller the price discount (price margin) between purchase and sale, the more favorable production is for yearling stocker operators. Of the ten years considered in setting expected beef prices, 1979 was the worst price year (widest price margin) and 1980, 1986, and 1987 were the best price years for yearling production.

Other model parameters. Using cost and return estimates prepared for yearling stocker operators in Kansas (McReynolds and Barnaby) and in New Mexico (Torell, Williams, and Brockman), variable production expenses (r),

including salt and minerals, fuel and repairs, veterinary and medicine, labor, marketing, and miscellaneous expenses, were estimated to be \$55 per head over the 150-day grazing season. Fixed production expenses (a) were estimated to be \$18 per hectare including a return to operator labor and management, such that the objective function value represents a discounted return to land and improvements. The return stream was discounted at a 7% rate of interest (ρ). Other computer runs were made to evaluate the sensitivity of the solution to the discount rate.

To analyze differences in the time path of rangeland productivity under both the single-period and dynamic models, a forty-year planning period with two years of deferment ($t_0 = 2$) was considered. Treatment cost (K) was \$41 per hectare, the average cost for control of sagebrush using tebuthiuron pellets (McDaniel et al.). Additional computer runs were also made to see how cost of treatment affected the optimal re-treatment strategy.

Application Results

Optimal stocking rates. The optimal stocking rate varies from year to year (fig. 1). In good price years, with a relatively small price discount for heavier cattle, like 1980, 1986, and 1987 (table 3), the optimal stocking rate is nearly twice that of unfavorable price years like 1979 when the price discount is large. Standard stocking rate recommendations often recognize that grazing use rates should vary from year to year, de-

⁷ With a positive discount rate, numerical solution could not be obtained for over a 40-year grazing rotation; thus, this was the maximum rotation considered.

⁸ In the actual analysis, purchase weight was constant at 215 kg, but sale weight varied depending on the optimal stocking rate decision.

Table 3. Beef Prices Used in the Economic Analysis

Price Year	Year in the Analysis	Futures Price		Estimated Price		Price Discount
		Spring	Fall	Purchase	Sale	
----- (\$/kg) -----						
1979	1, 11, 21, 31	2.01	1.83	1.79	1.44	-0.35
1980	2, 12, 22, 32	1.50	1.65	1.55	1.39	-0.16
1981	3, 13, 23, 33	1.57	1.46	1.59	1.32	-0.27
1982	4, 14, 24, 34	1.48	1.46	1.54	1.33	-0.21
1983	5, 15, 25, 35	1.43	1.30	1.52	1.27	-0.25
1984	6, 16, 26, 36	1.41	1.43	1.51	1.32	-0.19
1985	7, 17, 27, 37	1.43	1.39	1.52	1.30	-0.22
1986	8, 18, 28, 38	1.17	1.37	1.40	1.30	-0.10
1987	9, 19, 29, 39	1.54	1.73	1.57	1.41	-0.16
1988	10, 20, 30, 40	1.74	1.70	1.67	1.40	-0.27

Note: Specification of the Schroeder et al. price model was with a purchase weight of 215 kg, variable sale weight depending on the stocking rate decision, 50-head lot size, healthy fleshy condition, mixed horns and breeds, shrunk fill, medium muscling, medium upper and lower frame, purchased during 2nd quarter at market 5, sold during 4th quarter at market 5 and with future prices (\$/kg) as specified above.

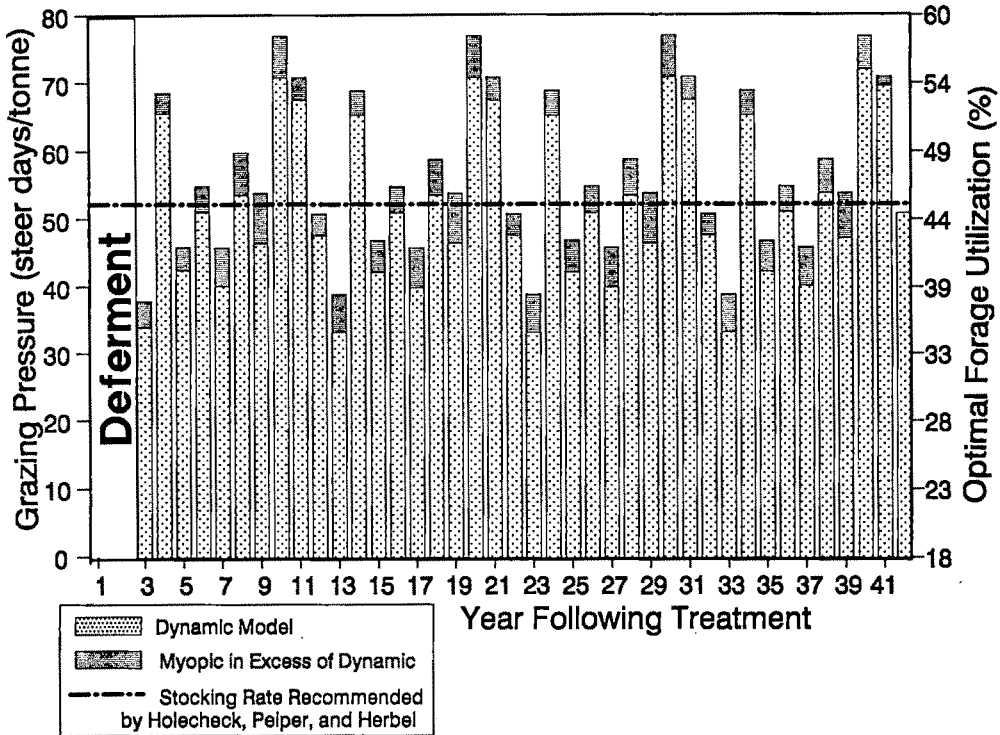


Figure 1. Time path of optimal grazing pressure and forage utilization

pending on forage production, but variation due to economic conditions is rarely recognized.⁹

In most years, economically optimal *GP* would be near that realized under the moderate stocking rate definition of the Colorado grazing study (table 1). Defined heavy use rates would be optimal during good price years.

The current stocking rate recommendation for this midgrass prairie site is to utilize 45% of available forage (Holechek, Pieper, and Herbel, p. 190). This implies a constant *GP* of fifty-two steer days per metric ton of forage. The recommendation, based largely on noneconomic factors, is remarkably close to the average economic optimal use rate shown in figure 1. Adjusting stocking rates to economic conditions is important, however. By adjusting stocking rates up and down depending on economic conditions, the objective function value was increased by 24%, from minus \$63 per hectare with the

constant 45% utilization rule to minus \$48 per hectare with flexible, economically optimal stocking rates.

Optimal stocking rates determined by the dynamic and single-period (myopic) models were not greatly different. As expected, when impacts to future forage production were included in the stocking rate decision, economically optimal stocking rates were decreased, but only slightly. This implies, for this midgrass prairie range site, that the impact of grazing pressure on current period animal performance drives the economic stocking rate decision. This is different than the results reported by Pope and McBryde, where diminished rangeland productivity drove the economic stocking rate decision.

As the planning period draws to a close, the difference between optimal stocking rates defined by the dynamic and myopic stocking rate models is reduced until the final year, when the two rates are equal. This occurs because there is no additional future value to remaining forage; i.e., the opportunity cost of increased grazing is zero because the grass stand will be rejuvenated next year.

⁹ Holechek, Pieper, and Herbel (p. 186) recognize four basic strategies that producers might select: (1) stock conservatively so that the range is not overstocked, (2) stock at a high rate commensurate with forage availability during favorable years, (3) stock at the average, and (4) vary stocking to meet forage supplies for each year.

Optimal rangeland productivity. Corresponding to the fluctuations in optimal stocking rates, average herbage production varies optimally through time after reaching a long-term level of production equilibrium (fig. 2), similar to that found under moderate use rates in the Colorado grazing study (table 1). Further, as shown with other applications in natural resource economics (Samuelson, Fisher), producing at a maximum sustainable yield would not be economically optimal. Peak forage production was not maintained under profit-maximizing stocking rates with either the myopic or dynamic models. Rangeland productivity was not devastated either. After an initial drop in average herbage production, optimal rangeland production declined to about 1,500 kilograms per hectare under the dynamic model and 1,450 kilograms per hectare for the myopic model. Implications are that maintaining maximum sustainable grass yield and range condition, a primary goal in range management, is unobtainable, given the profit motive of ranchers. But, for this range site, the profit motive will not ruin the range either.

Net present value. By reducing stocking rates and foregoing revenue in the early years of the planning period, revenue in future years was increased only slightly for the dynamic model ($-\$43/\text{ha}$ for the dynamic model versus $-\$49.33/\text{ha}$ for the myopic model). As indicated by the negative objective function values, alternative resource uses may be expected over the long term. Annual returns to land and improvements were only positive with the 1980, 1984, 1986, and 1987 beef price situations. Variable production expenses were covered in all years. Price margins between yearling purchase and sale have widened in recent years, and

the profit potential for yearling stocker operations on rangeland, relative to cow/calf and sheep ranches, has diminished (Torell, Williams, and Brockman).

Optimal retreatment schedule. Given a $\$41$ per hectare treatment cost, plus two years of grazing deferment, there is no need for retreatment on the Colorado study site. Model results indicate grass productivity would be maintained by grazing at economically optimal stocking rates. Realistic beef prices faced by livestock producers will cause economically optimal stocking rates to lie below levels that would significantly deteriorate rangeland productivity. In this application, no additional investment through significantly reduced grazing use rates or rangeland brush control treatments is needed or economically justified.

By extending the grass rotation from every seventeen years to every twenty-two years, the value of the objective function increased from $-\$73$ per hectare to $-\$58$ per hectare. Increasing the rotation by an additional ten years (32-year rotation) decreased the loss still further to $-\$51$ per hectare. With forty years of grazing use before retreatment, the objective function value was reduced slightly more to $-\$48$ per hectare. If a rancher started with the Colorado range site already at peak forage production, such that no initial treatment and deferment were necessary, managed the site over an infinite planning horizon, and used economically optimal stocking rates every year, the specified costs and prices considered in the analysis would indicate a nearly breakeven investment. In this case, the objective function value (return to land and improvements) would be $-\$5$ per hectare.

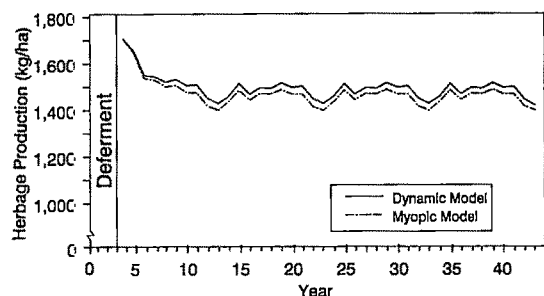


Figure 2. Time path of optimal herbage production

Sensitivity analyses. As shown in figure 1, optimal GP was sensitive to changes in expected beef prices and annual production costs. This was true of both the dynamic and myopic models. Additional computer runs were made to see how sensitive the optimal solution was to rangeland treatment costs (K) and discount rate (ρ). In general, the solution was not significantly affected by changes in either of these parameters.

As shown by Perrin, increasing the discount rate generally results in an earlier optimum replacement age, as occurred in this application. Because the model would not solve for more than

a forty-year planning period, we could not establish an exact optimal time for replacement. Empirical results indicated, however, as the discount rate increased from 7% to the unrealistic level of 30%, that the optimal grazing use rate increased by about two steer days per metric ton of forage, implying a faster use rate and more rapid replacement. This was an insignificant difference because a two-unit change in *GP* would not even be measurable.

The value of the objective function was increased with an increase in discount rate. Increasing the discount factor from 7% to 14% increased the objective function value from -\$48 per hectare to -\$44.91 per hectare because future losses became less important in present value terms.

Optimal stocking rates were not affected by the cost of the rangeland improvement. Similarly, the optimal retreatment schedule was insensitive to treatment cost, except at unrealistically low levels. When the treatment cost was greatly reduced, from \$41 per hectare to less than \$5 per hectare, the optimal replacement cycle was every four years (plus the 2 years of deferment). Above this cost level, optimal replacement stayed at the forty-year cycle found with the base run. This is because little would be gained from implementing a range improvement as implied by the equation system defined in the Colorado application. As shown in figure 2, by practicing economically optimal stocking rates, herbage production would only decline by about 200 kilograms per hectare, and this would occur within five years.

Model Limitations

The dynamic stocking rate model developed here is deterministic with the assumption that at the time the stocking rate decision is made a rancher can accurately predict prices, costs, and forage conditions. Stochastic weather variations are not directly considered in the stocking rate decision; the time path of average rangeland productivity defines the intertemporal dynamics of the model. This procedure is appropriate for the yearling stocker operations considered in the model application. This type of livestock operation has a considerable marketing flexibility and can annually adjust the number of head purchased, depending on both economic and forage conditions. Our results show that a yearling operator should carefully consider expected economic and

forage conditions for the year and set stocking rates accordingly.¹⁰ Furthermore, the results indicate that little if any thought need be given to what the stocking rate decision will do to forage production in future years.

The stocking rate decision is more complicated when brood stock must be maintained between years. In this case, future weather patterns are an important consideration. A major cost will be incurred in selling and buying brood stock or in purchasing supplemental feed if stocking rates are set too low or high relative to realized forage production. Unlike the yearling stocker operator, cow/calf and sheep ranchers must make long-term stocking rate decisions. In this case, planning stocking rates based on the average forage production expected under alternative stocking rates may not be appropriate.

Concluding Comments

As noted by Bromley, an important grazing interaction is the impact of grazing on future rangeland productivity. This is universally believed, as evident by the rationale given for recommending desirable stocking rates (Holechek, Pieper, and Herbel; Stoddart, Smith, and Box). However, results of this study indicate that intertemporal grazing impacts to forage production may not be that important after all. Rather, our results show that dependence of current period animal performance on stocking rate is the key. Further, the model results show that ranchers have no economic incentive as profit maximizers to overgraze continually. We found, for a Colorado study site, with realistic expectations of beef prices and production costs, falling animal performance will set profit-maximizing stocking rates well below those levels that will severely deteriorate the rangeland resource. Similarly, Torell also found current period animal production impacts to limit the motivation to overgraze the range. Thus, like Workman (p. 52), we conclude that historic overgrazing can be blamed on ignorance and overoptimism about rangeland carrying capacities, but this overgrazing occurs in spite of the profit motive not because of it.

Our results, (a) animal performance driving the economic stocking rate decision and (b) no

¹⁰ If warm-season grasses that mature late in the season predominate the range site, then the uncertainty about forage production this year will be an important consideration (Rodriguez and Taylor).

economic incentive to continually overgraze the range, are not universally true. As a counter example, Pope and McBryde developed a dynamic stocking rate model for cow/calf producers in south Texas. They conclude, "if economically viable and ecologically sound range improvement treatments are available, a range management strategy that systematically overstocks and periodically applies treatments would be more profitable" (p. 167). Either result may be valid, depending on type of livestock grazed, the cost of feasible rangeland improvements, and the range site under consideration.

The idea that animal performance impacts from grazing can be the most important consideration in stocking rate decisions is a new idea for most people. The implications for management and research are many. Most important, if dynamic impacts of grazing to rangeland productivity are small relative to current period production impacts, this implies the standard economic model of optimal input use can satisfactorily be used to make nearly optimal stocking rate decisions without the more complicated modeling of dynamic interactions. More sophisticated models are not always needed.

[Received February 1990; final revision received October 1990.]

References

- Bransby, D. I., B. E. Conrad, H. M. Dicks, and J. W. Drane. "Justification for Grazing Intensity Experiments: Analyzing and Interpreting Grazing Data." *J. Range Manage.* 41(1988):274-79.
- Bromley, D. W. "A Dynamic Economic Model of Pasture and Range Investments: Comment." *Amer. J. Agr. Econ.* 54(1972):131.
- Burt, O. R. "A Dynamic Economic Model of Pasture and Range Improvements." *Amer. J. Agr. Econ.* 53(1971):197-205.
- . "A Dynamic Economic Model of Pasture and Range Investments: Reply." *Amer. J. Agr. Econ.* 54(1972):131-32.
- Chavas, J.-P., J. Kliebenstein, and T. D. Crenshaw. "Modeling Dynamic Agricultural Production Response: the Case of Swine Production." *Amer. J. Agr. Econ.* 67(1985):636-46.
- Fisher, A. C. *Resource and Environmental Economics*. New York: Cambridge University Press, 1981.
- Hardin, G. "The Tragedy of the Commons." *Science* 162(1968):1243-48.
- Hart, R. H. "Economic Analysis of Stocking Rates and Grazing Systems." *Proceedings, The Range Beef Cow Symposium X, A Symposium on Production*. Cheyenne WY, 8-10 Dec. 1987. University of Wyoming Coop. Extens. Serv.
- Hart, R. H., M. J. Samuel, P. S. Test, and M. A. Smith. "Cattle, Vegetation, and Economic Responses to Grazing Systems and Grazing Pressure." *J. Range Manage.* 41(1988a):282-86.
- Hart, R. H., J. W. Waggoner, Jr., T. G. Dunn, C. C. Kaltenbach, and L. D. Adams. "Optimal Stocking Rate for Cow-calf Enterprises on Native Range and Complementary Improved Pasture." *J. Range Manage.* 41(1988b):435-41.
- Hildreth, R. J., and M. E. Riewe. "Grazing Production Curves II: Determining the Economic Optimum Stocking Rate." *Agron. J.* 55(1963):370-72.
- Holechek, J. L., R. D. Pieper, and C. H. Herbel. *Range Management Principles and Practices*. Englewood Cliffs NJ: Prentice-Hall, 1989.
- Karp, L., and A. Pope, III. "Range Management Under Uncertainty." *Amer. J. Agr. Econ.* 66(1984):437-46.
- Long, N. V., and N. Voutsden. "Optimal Control Theorems." *Applications of Control Theory to Economic Analysis*, ed. J. D. Pitchford and S. J. Turnovsky. New York: North-Holland Publishing Co., 1977.
- McDaniel, K. C., L. A. Torell, J. M. Fowler, and K. W. Duncan. "Brush Control on New Mexico Rangeland." New Mexico State University Coop. Extens. Serv. Guide 400-B-18, 1986.
- McReynolds, K. L., and G. A. Barnaby, Jr. *Grazing Yearling Beef*. Kansas State University Cooperative Extens. Serv. Farm Management Guide MF-591, 1986.
- Murtaugh, B. A., and M. A. Saunders. "A Projected Lagrangean Algorithm and Its Implementation for Sparse Nonlinear Constraints." Dep. Operations Res. Tech. Rep. SOL-80-1R, Stanford University, 1981.
- Perrin, R. K. "Asset Replacement Principles." *Amer. J. Agr. Econ.* 54(1972):60-67.
- Pope, C. A., and G. L. McBryde. "Optimal Stocking of Rangeland for Livestock Production within a Dynamic Framework." *West. J. Agr. Econ.* 9(1984):160-69.
- Riggs, W. W. "Economic Principles for Myopic Versus Dynamic Stocking Rate Decisions." M.S. thesis, New Mexico State University, 1989.
- Rodriguez, A., and R. G. Taylor. "Stochastic Modeling of Short-Term Cattle Operations." *Amer. J. Agr. Econ.* 70(1988):121-32.
- Riewe, M. E. "The Economics of Grazing." *Forage Evaluation: Concepts and Techniques*, ed. J. L. Wheeler and R. D. Mochrie, pp. 517-26. (Proceedings of Bilateral Workshop held in Armidale, NSW, Australia, 27-31 Oct. 1980.) Lexington, KY: American Forage and Grassland Council, 1981.
- Samuelson, P. A. "Economics of Forestry in an Evolving Society." *Econ. Inquiry* 14(1976):466-92.
- Schroeder, T., J. Mintert, F. Brazle, and O. Grunewald. "Factors Affecting Feeder Cattle Price Differentials." *West. J. Agr. Econ.* 13(1988):71-81.
- Sims, P. L., B. E. Dahl, and A. H. Denham. "Vegetation and Livestock Response at Three Grazing Intensities on Sandhill Rangeland in Eastern Colorado." Colorado

- State University Agr. Exp. Sta. Tech. Bull. No. 130, 1976.
- Stevens, J. B., and E. B. Godfrey. "Use Rates, Resource Flows, and Efficiency of Public Investment in Range Improvements." *Amer. J. Agr. Econ.* 54(1972):611-21.
- Stoddart, L. A., A. D. Smith, and T. W. Box. *Range Management*. New York: McGraw-Hill Book Co., 1975.
- Torell, L. A. "Economic Optimum Stocking Rates and Re-treatment Schedule for Crested Wheatgrass Stands." Ph.D. thesis, Utah State University, 1984.
- Torell, L. A., and R. H. Hart. "Economic Consideration for Efficient Stocking Rates on Rangeland." *Achieving Efficient Use of Rangeland Resources*, ed. R. S. White and R. E. Short, p. 71-76. Proceedings of Rangeland Use Symposium, Montana State University, 1988.
- Torell, L. A., A. Williams, and B. A. Brockman. *Live-stock Cost and Return Estimates for New Mexico, 1986*. New Mexico State University Agr. Exp. Sta. Res. Rep., 1989.
- Vallentine, J. F. *Grazing Management*. San Diego: Academic Press, 1990.
- Wight, J. R., ed. *SPUR-Simulation of Production and Utilization of Rangelands: Documentation and Users Guide*. Washington DC: U.S. Department of Agriculture, ARS Publication No. ARS-63.
- Workman, J. P. *Range Economics*. New York: Macmillan Co., 1986.

Testing for Consistent Aggregation

Robert G. Chambers and Rulon D. Pope

This paper designs workable empirical procedures for testing or imposing supply (demand)-price aggregators. Only single-equation methods are considered to shift the focus from any particular optimizing paradigm toward the necessities for correctly tabulating firm-level supplies (demands) and aggregate supply (demand). Both macro- and micro-level tests for consistent aggregation are developed. And constructive rules for aggregation based upon empirically observed micro behavior are provided in tabular form. Empirical tests reveal no support for the aggregate price most commonly reported by USDA. The constructive rules are then used to develop a new wheat-price index consistent with supply aggregation.

Key words: aggregate demand, aggregate supply, exact aggregation, price indexes.

A substantial literature has evolved on testing for the existence and form of quantity or price aggregates (e.g., Weaver, Shumway, Ball, Chavas and Cox). These tests generally rely on separability restrictions and aggregates are formed (horizontally) either across distinct commodities (livestock, grains, etc.) or commodity prices. Dual tests based on profit, cost, or expenditure functions search for the existence of aggregate prices, while primal tests search for the existence of quantity aggregates. Under suitably strong assumptions (weak homothetic separability) dual and primal tests are equivalent.

A closely related body of literature focuses on the construction of aggregates (vertically) over individual consumers and producers (e.g., Muellbauer; Nichols; Gorman 1953, 1968; Stoker; Jorgenson, Lau, and Stoker; Antle; Chambers). This literature's motivation typically is the construction of an aggregate income, output, or capital index to be used in aggregate demand, cost, or profit systems. This approach, usually referred to as exact aggregation, is nonparametric in a statistical sense because it aggregates any income distribution. Typical of the questions asked by this approach is: "If per capita income is used in aggregate demand functions and if aggregate demand is the sum of individual demands, what form must the micro and macro demands take?" (Gorman 1953). The Gorman

polar form (GPF) is the answer. The almost ideal demand system (AIDS), a special case of Muellbauer's PIGLOG system, also aggregates consistently using Muellbauer's approach to aggregation. Numerous papers have either tested for or assumed exact aggregation using the Muellbauer approach (Deaton and Muellbauer; Nichols; Jorgenson, Lau, and Stoker; Blanciforti and Green; Eales and Unnevehr; Green and Alston).

Another empirical reality confronting agricultural economists is that not only incomes but prices also typically vary across economic agents. Heterogenous or disperse prices, while hard to explain in terms of simple neoclassical theory, are a fact of life. They can occur because of quality, timing, and spatial differences, government programs (e.g., nonrecourse loan rates are specified on a county basis), regulations, and uncertainty (heterogenous expectations). As just one example, empirical analysis consistently reveals that state-level prices of land, labor, and outputs are not nearly colinear across states (Robison, Lins, and VenKataraman; Gould and Saupe).

All aggregate supply and demand functions must inevitably deal with this reality even though to some it may contradict the usual predictions of perfect competition. The price of wheat, for example, can differ across individuals or other economic entities because of all of the above factors and still be consistent with price-taking behavior by wheat growers. To be sure, prices of different wheat "qualities" at different points in space or time are not entirely independent.

Robert G. Chambers is a professor, Department of Agricultural and Resource Economics, University of Maryland; Rulon D. Pope is a professor, Department of Economics, Brigham Young University.

But, just as certainly, examination of the data and the well-known fact of basis variations reveals that correlations across these different qualities are not nearly one. For these reasons, U.S. Department of Agriculture (USDA) price indexes must (and do) aggregate spatially and intertemporally in order to report the U.S. annual average price of a commodity. Nowhere is this more apparent than in the season average prices reported in the USDA publication *Agricultural Statistics* which weight state prices by quantities sold and include allowances for unredeemed loans and government purchases at state average loan rates (which, if applicable, are administered and not competitive prices).

This paper designs workable empirical procedures for testing or imposing supply-price aggregators which capture accurately the information that aggregate analyses are trying to uncover. Unlike Deaton and Muellbauer, only single-equation problems are studied. This specification shifts the focus away from any particular optimizing paradigm toward the necessities of correctly tabulating firm-level demands and supplies. Thus, the approach is robust to various optimization paradigms. But the single-equation approach is also adopted for two other reasons. First, it avoids the negativism associated with aggregation of systems (Pope and Chambers). Second, most applied work in agricultural economics does not employ systems-type restrictions; therefore, the methods and results here should have wide applicability.

The approach enables researchers to delineate the class of supply or demand functions consistent with the actual aggregate price indexes employed. One can then test for consistent aggregation by testing whether the implied form is appropriate. Both macro and micro tests are developed. Although some of the analysis is quite general, particular attention is paid to the two most commonly published producer price indexes, the simple and weighted averages.

In what follows, we first motivate the need for testing by illustrating the effects of using a Cobb-Douglas supply or demand function and either of the two most commonly published price indexes. The third section generalizes the above examples and provides constructive rules for consistent aggregation. The fourth section develops and applies empirical tests for consistent aggregation. Significantly, we find no empirical support for the state-level technology implied by presently calculated U.S. season average wheat prices. We propose and calculate an alternative

price index which is consistent with aggregation. The fifth section concludes. Although for illustrative purposes our discussion is solely in terms of producer models, the results apply directly to the consumer case.

Common Agricultural Price Aggregation and the Cobb-Douglas: An Example

The most common aggregate prices used by agricultural economists are simple averages or weighted averages. Prices reported in such familiar sources as *Agricultural Statistics*, *Agricultural Prices*, or the *Survey of Current Business* tend to be measured in this way. For example, *Agricultural Prices* reports commodity prices received by farmers by states on a monthly basis. These prices usually are "estimates of average prices received for all of the commodity sold during the entire month." Thus, the price-aggregation procedure is simple averaging. This same publication, as well as *Agricultural Statistics*, also reports annual U.S. season average prices for many commodities. These U.S. season average prices are computed by weighting "state season average prices by the estimated quantity sold in each state." Here the aggregation procedure is weighted averaging. (We refer to this particular weighted average as the Laspeyres in what follows.) This section briefly considers the implications of using these two price aggregation schemes in Cobb-Douglas supply-response models.

Consider the price-aggregation problem: The goal is to relate aggregate supply, Y (the sum of firm supplies), econometrically to the average industry price, $\bar{P} = 1/m \sum_{j=1}^m p_j$ (p_j is the price facing the j th firm) and a vector of input prices, w , which are assumed to be constant across all firms.

To do so, the researcher first specifies an aggregate supply function which we denote by $G(\bar{P}, w)$. Then using data on Y , \bar{P} , and w , the parameters of G are estimated. For many agricultural applications, G might be chosen as Cobb-Douglas, $G \equiv A\bar{P} w^{-a}$. For notational simplicity, assume w is a scalar.

A basic question is: Is this procedure consistent in the sense that G accurately represents the sum of firm supplies over all m firms? Differentiate both sides of the aggregate supply relationship ($Y = G(\bar{P}, w)$) with respect to p_i and p_j ($j \neq i$) using the definition of Y , G , and \bar{P} to

find that a necessary condition for a positive answer is

$$(1) \quad \frac{\frac{\partial G}{\partial \bar{P}} \frac{\partial \bar{P}}{\partial p_i}}{\frac{\partial G}{\partial \bar{P}} \frac{\partial \bar{P}}{\partial p_j}} = \frac{\frac{\partial \bar{P}}{\partial p_i}}{\frac{\partial \bar{P}}{\partial p_j}} = \frac{\partial Y / \partial p_i}{\partial Y / \partial p_j} = \frac{\partial y_i / \partial p_i}{\partial y_j / \partial p_j} \quad \text{for all } i, j = 1, \dots, m.$$

Because \bar{P} is the simple average, expression (1) reduces to

$$\frac{\frac{\partial \bar{P}}{\partial p_i}}{\frac{\partial \bar{P}}{\partial p_j}} = 1 = \frac{\partial y_i / \partial p_i}{\partial y_j / \partial p_j} \quad \text{for all } i, j = 1, \dots, m.$$

Supply slopes must be equal for all firms for aggregation to be consistent. Because each such slope can depend only upon p_i and w (firm supply depends only on the prices it faces), these slopes must be independent of the firm and hence depend only on w . Integrating with respect to p_i yields

$$y_i = \alpha_i(w) + \beta(w)p_i \quad i = 1, \dots, m.$$

That is, firm-level supplies must be linear in own price. Rejection of this form implies that mean price cannot consistently aggregate supplies. Summing shows that aggregate supply is also linear:

$$Y \equiv \sum_{i=1}^m y_i = \sum_{i=1}^m \alpha_i(w) + \beta(w) \sum_{i=1}^m p_i = \alpha(w) + \beta(w) \cdot m\bar{P},$$

where $\sum_{i=1}^m \alpha_i(w) \equiv \alpha(w)$. Micro and macro supplies must satisfy these functional relations if consistent aggregation is to be possible.

Suppose for the moment that micro supplies are indeed aggregable and then reveal that G has been chosen as Cobb-Douglas so that

$$(2) \quad Y = A\bar{P}^a w^{-a} = \alpha(w) + \beta(w)m\bar{P}.$$

This implies $Aw^{-a}\bar{P}^a = \beta(w)m\bar{P}$ and $\alpha(w) = 0$. A regression with this commonly used form must tend toward a unit elasticity supply curve as the aggregation errors tend toward zero. The aggregate supply elasticity must be unity! This conclusion can also serve as the basis for testing exact aggregation with mean price using the aggregate supply function. This latter equality also

requires that either $\alpha_i(w) = 0$ ($i = 1, \dots, m$) or that $\sum_{i=1}^m \alpha_i(w) = 0$ with some $\alpha_i \neq 0$. In the case that all $\alpha_i(w) = 0$, micro and macro supply elasticities are unity.

Now, consider the Laspeyres price index

$$P = \hat{P} = \frac{\sum_{i=1}^m p_i y_i}{\sum_{i=1}^m y_i}.$$

Again assume that $\sum_{i=1}^m y_i \equiv Y$. \hat{P} clearly is a function of Y and w . Because the denominator of \hat{P} is Y , it seems clear that the numerator depends on Y and w . Pope and Chambers have shown that the numerator (revenue) of \hat{P} is necessarily affine in Y . Thus, $\sum_{i=1}^m p_i y_i = -\alpha(w) + \beta(w)Y$ with $p_i y_i = -\alpha_i(w) + \beta(w)y_i$. Hence,

$$\hat{P} = \frac{-\alpha(w) + \beta(w)Y}{Y} \quad \text{and}$$

$$Y = \frac{\alpha(w)}{\beta(w) - \hat{P}}.$$

Only forms satisfying this last equation are candidates for $G(\hat{P}, w)$. For example, suppose as above that $G(\hat{P}, w)$ is Cobb-Douglas $A\hat{P}^a w^{-a}$. Comparing this with the above then requires $\beta(w) = 0$, $A = 1$, $a = -1$. Hence, a Cobb-Douglas can be used with the Laspeyres only if aggregate supply has a supply elasticity of minus one! Aggregate and micro revenues are constant. This condition forms the basis of a test for consistent aggregation with the Cobb-Douglas macro function and Laspeyres index. It also raises the possibility that the index and functional form may unwittingly tend to prescribe empirical work.

Single-Equation Price Aggregation-General Results

The previous examples are important special cases of a more general price aggregation problem: Find functions G , P , and H satisfying

$$(3) \quad G(P, w) = H(h_1(p_1, w), h_2(p_2, w), \dots, h_m(p_m, w)) \\ P = P(p_1, \dots, p_m; w).$$

Here p_i remains the price faced by the i th firm, and w is the vector of input and output prices which are common to all firms. The indexes ($i = 1, \dots, m$) delimit the number of firms

whose supplies or demands are aggregated. The functions h_i are the microeconomic response functions (e.g., supplies, demands, profits, or costs) to be aggregated. The function H represents the rule by which the aggregate quantity variable is constructed from these microeconomic response functions. For example, in both examples above, h_i is the supply of the i th firm and H is the summation operator. The function G is the aggregate behavioral function upon which analysis is based. In the examples above, G is the Cobb-Douglas industry supply function; P is the representative price used in the aggregate relationship. In the first example, P is the average price, while in the second it is the weighted average. Aggregation schemes which satisfy (3) are said to aggregate consistently.

Attention is restricted to strictly monotonic and differentiable H functions and G functions. This restriction represents only a slight loss in generality because virtually all commonly used functional forms satisfy these criteria. This assumption clarifies the intimate relationship between the aggregate price index and the microeconomic response functions.

Result 1. The aggregate price index P , consistent with (3), can always be expressed as a differentiable and monotonic function of the microeconomic response functions of the following form:

$$(4) \quad P = P^*(h_1(p_1, w), \dots, h_m(p_m, w); w).$$

P^* has partial derivatives

$$(5) \quad \frac{\partial P^*}{\partial h_i} = \left(\frac{\partial G}{\partial P} \right)^{-1} \left(\frac{\partial H}{\partial h_i} \right) \text{ and} \\ \frac{\partial P^*}{\partial w} \Big|_{h_i} = \left(\frac{\partial G}{\partial P} \right)^{-1} \left(\frac{\partial G}{\partial w} \right),$$

where the notation $\partial P^* / \partial w|_{h_i}$ denotes the partial derivative of P^* with respect to w holding the h_i functions constant.

Expressions (4) and (5) are direct consequences of the implicit function theorem (Courant, p. 117) applied to (3). The price-aggregation rule, P , inherits structural restrictions placed upon H . For example, suppose that the aggregate economic quantity is firmwise strongly separable in the microeconomic response functions. This includes the case where aggregate supplies, demands, shares, or profits are the sum of their micro counterparts. It also includes the case

where the log of the sum of the micro entities equals the log of the macro entity and so on. Functionally,

$$(6) \quad H(h_1, \dots, h_m) = \bar{H} \left(\sum_{i=1}^m g_i(h_i(p_i, w)) \right) \equiv \bar{H}(g).$$

Applying (4) to (6) gives the price aggregation rule consistent with (6),

$$(7) \quad P = \bar{P}^* \left(\sum_{i=1}^m g_i(h_i(p_i, w)); w \right) \equiv \bar{P}^*(g; w).$$

Expression (7) is firmwise strongly separable in the prices p_i . The reader can easily verify that expression (4) implies the following result (proved formally in Pope and Chambers):

Result 2. Let N be the set of integers $\{1, \dots, m\}$ and let $\Omega = (\Omega_1, \dots, \Omega_s, \dots, \Omega_s)$ be a mutually exclusive and exhaustive partition of N . Let Ω_p and Ω_h be the set of prices, p , and functions, h , partitioned according to the indexes in Ω . Then, for H , P , and h , monotonic and twice differentiable, with $H(h_1(p_1, w) \dots, h_n(p_n, w)) = G(P, w)$, P is strongly (weakly) separable in Ω if and only if H is strongly (weakly) separable in Ω .

Thus, one cannot divorce the structure of the price aggregate from the structure of the quantity aggregation rule. Once the researcher picks a price aggregate (quantity aggregation rule) the quantity aggregation rule (the price aggregate) cannot be chosen arbitrarily. Result 2 shows, for example, that choosing the price aggregate to be mean price (this rule is strongly separable) automatically rules out using quantity aggregation rules which are not strongly separable. Because most aggregation rules used in the construction of reported quantity aggregates are strongly separable aggregation rules, we consider further the practical implication of result 2.

Constructing Firmwise Strongly Separable Aggregates

To examine the implications of result 2, imagine that one chooses a price-aggregation rule

$$(8) \quad P = \bar{P} \left(\sum_{i=1}^m k_i(p_i, w); w \right) = \bar{P}(k; w)$$

which is firmwise strongly separable in the prices

p . [The average price is a special case of (8).] Result 2 now implies H must be also firmwise strongly separable in the h functions, i.e., it can be expressed as (6), to satisfy consistent aggregation. [Alternatively, by result 2, if one chooses a quantity aggregation rule satisfying (6), the price aggregation rule must satisfy (8).] Use this fact and differentiate both sides of (3) to determine in this case that

$$\frac{\partial G}{\partial p_i} = \frac{\partial G}{\partial P} \frac{\partial \bar{P}}{\partial k} \frac{\partial k_i}{\partial p_i} = \frac{\partial \bar{H}}{\partial g} \frac{\partial g_i}{\partial p_i} \quad i = 1, \dots, m.$$

Taking ratios gives, after a slight manipulation,

$$(9) \quad \frac{\partial k_i / \partial p_i}{\partial g_i / \partial p_i} = \frac{\partial k_j / \partial p_j}{\partial g_j / \partial p_j} \quad \text{for all } i \text{ and } j.$$

Expression (9) implies that these ratios must be the same for all i and j . By construction, each ratio can only depend upon the vector w and the price specific to the i th firm. Because they must be the same regardless of the firm considered (and thus the p_i), these ratios can only depend upon w —call this common ratio $\gamma(w)$. Integrating the differential equations implied by (9) gives

$$(10) \quad g_i(h_i(p_i, w)) = \gamma(w)k_i(p_i, w) + v_i(w) \quad i = 1, \dots, m$$

and

$$h_i(p_i, w) = g_i^{-1}[\gamma(w)k_i(p_i, w) + v_i(w)] \quad i = 1, \dots, m.$$

To aggregate consistently, the i th response function must be a monotonic transformation of an affine transformation of $k_i(p_i, w)$.

Limitations on Strongly Separable Aggregates

Table 1 presents a guide to constructing firmwise strongly separable price indexes and quantity indexes which are mutually consistent. The families of price indexes in the second column exhaust the possible alternatives consistent with the quantity indexes in the corresponding row of the first column. If the researcher chooses a price aggregate belonging to the second column of that table, the structure of the associated micro response functions is predetermined. Notice, in particular, that this relationship does not follow from any optimization hypothesis but from the arithmetic necessity of having both sides of aggregate equations consistent with one an-

other. Violating these principles in empirical work is, to use an old saw, "adding apples and oranges." Just as two apples and two oranges do not make four apples or four oranges, using inconsistent aggregation procedures violates the simple arithmetic laws upon which aggregation is purportedly based.

The limiting nature of these results can be seen clearly by considering using a geometric average of firm-level prices as a price index (row 5, column 2) to explain industry supply (the sum of firm supplies). From table 1, the firm-level supply functions consistent with aggregation must belong to the family of functions,

$$\alpha_i(w) + \hat{\beta}(w)\theta_i(w)\ln p_i \quad i = 1, \dots, m.$$

Why Quantity (Price) Aggregates Inherit the Structure of Price (Quantity) Aggregates

When either the price or the quantity index is specified to be firmwise strongly separable, its counterpart in the aggregation scheme inherits both this structural property as well as the general functional structure imposed. Why is this result obtained? The key lies in recognizing that the three level sets associated with the aggregate response function (G), the aggregate price index (P), and the aggregate quantity index (H), respectively

$$\{p: G(P(p_1, \dots, p_m); w) = G\}$$

$$\{p: P(p_1, \dots, p_m; w) = P\}$$

$$\{p: H(h_1(p_1, w), \dots, h_m(p_m, w)) = H\}$$

have the same graph in p -space for fixed w . Thus, once one specifies the curvature properties of, say, the price aggregation rule in p -space, one also specifies the curvature of the level set of the quantity aggregation rule and the aggregate response function. Put another way, each of these functions when considered as functions of the prices over which one is aggregating are monotonic transformations of the others (see result 1). Hence, once one of these functions is specified (say H), the whole family of possible functions available for the two other functions (G and P) is known. Because these latter families correspond to monotonic transformations of the original aggregation rule specified (there is an infinity of such possible transformations), there is an infinity of possible rules which are consistent with aggregation over prices. This is reflected

Table 1. The Structure of Price Aggregators and Aggregation Functions

Aggregate Function (<i>H</i>)	Class of Price Aggregators <i>P</i>
General structure	
$H\left(\sum_{i=1}^m h_i(p_i, w)\right)$	$P^*\left(\sum_{i=1}^m (\alpha_i(w) + \beta(w)h_i(p_i, w)); w\right)$
$H\left(\sum_{i=1}^m \hat{\alpha}_i(w) + \hat{\beta}(w)k_i(p_i, w)\right)$	$P^*\left(\sum_{i=1}^m k_i(p_i, w); w\right)$
Specific examples	
$\sum_{i=1}^m A_i p_i^\alpha w^{1-\alpha}$	$P^*\left(\sum_{i=1}^m \alpha_i(w) + \beta(w)A_i p_i^\alpha w^{1-\alpha}; w\right)$
$H\left(\sum_{i=1}^m \hat{\alpha}_i(w) + \frac{\hat{\beta}(w)}{m} p_i\right)$	$\frac{1}{m} \sum_{i=1}^m p_i$
$H\left(\sum_{i=1}^m \hat{\alpha}_i(w) + \hat{\beta}(w)\theta_i(w) \ln p_i\right)$	$\prod_{i=1}^m p_i^{\theta_i(w)} = \exp\left[\sum_{i=1}^m \theta_i(w) \ln p_i\right]$
$\sum_{i=1}^m (a_i p_i^{\alpha_i} + b_i w^{\alpha_i})^{1/\alpha_i}$	$P^*\left(\sum_{i=1}^m \alpha_i(w) + \beta(w)(a_i p_i^{\alpha_i} + b_i w^{\alpha_i})^{1/\alpha_i}; w\right)$
$H\left(\sum_{i=1}^m \hat{\alpha}_i(w) + \hat{\beta}(w)c_i p_i^{\alpha(w)}\right)$	$\left(\sum_{i=1}^m c_i p_i^{\alpha(w)}\right)^{1/\alpha(w)}$

in table 1, for example, by the fact that the average price equation will aggregate firm-level supplies of the general form for any possible *H* function that is monotonic in its argument.

However, once a specific *H* function is chosen, say the sum, then choice of a specific price aggregator places direct restrictions on the form of the functions to be aggregated. Only special classes of *h*() functions can then be consistent with both *H*, *P*, and the requirement that the latter be expressible as a monotonic transformation of the former. This is why, once exact forms of *P* and *H* are specified, we can infer exactly the type of supply or demand functions (if any exist) that are consistent with these aggregators. For example, the Laspeyres price index is not generally firmwise strongly separable. But $H = \sum_{i=1}^m h_i$ is firmwise strongly separable. Thus, if this *H* is to be used with a Laspeyres index, only firm-level supply functions which make the Laspeyres index firmwise strongly separable in the *p_i*'s are consistently aggregable. Pope and Chambers demonstrate that these functions all belong to the family

$$h_i(p_i, w) = \frac{\alpha_i(w)}{\beta(w) - p_i}.$$

Therefore, consistent aggregation using the Laspeyres index requires that supplies be of this microform. If this form is rejected by the data, so also is consistent aggregation using the index.

These ideas are portrayed geometrically in figure 1. The level curves for *G*, *P*, and *H* are represented by the curve (set) *a*–*b*. The marginal rate of substitution of the curve is $(\partial P / \partial p_j) / (\partial P / \partial p_i)$ or equivalently $(\partial h_j / \partial p_j) / (\partial h_i / \partial p_i)$ which is nonnegative. In general the marginal rate of substitution may rise or fall with increasing *p_i*. For strongly separable aggregation rules,

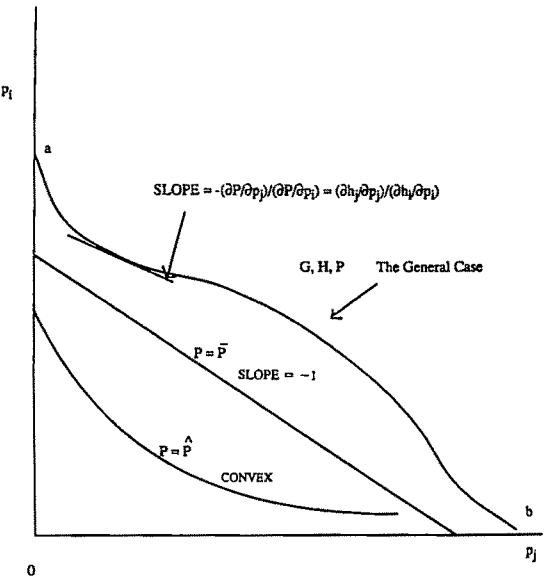


Figure 1. Level curves and aggregation

the marginal rate of substitution is unaffected by variations in p_k ($k \neq i, j$). When P equals average price, \bar{P} , the marginal rate of substitution is constant at unity. In the Laspeyres case, \bar{P} , the slope is nonconstant and equal to $\alpha_j(w)(\beta(w) - p_j)^{-2}/\alpha_i(w)(\beta(w) - p_i)^{-2}$. The marginal rate of substitution is diminishing, and hence aggregate price increases with a mean-preserving spread in the distribution of prices. Contrasting \bar{P} and P we see that mean-preserving changes in the p -distribution must have no effect on the aggregates if \bar{P} is used, but mean-preserving spreads in the p -distribution evoke aggregate change if P is used.

Finally, exact knowledge of P , H , and the h functions lets one infer $G(P, w)$ because differentiating (3) gives

$$(11) \quad \frac{\partial G}{\partial P} \frac{\partial P}{\partial p_i} = \frac{\partial H}{\partial h_i} \frac{\partial h_i}{\partial p_i}.$$

If $\partial P/\partial p_i$ and $\partial H/\partial h_i$ $\partial h_i/\partial p_i$ are known, (11) can be solved for $\partial G/\partial P$, which can, in principle, be integrated to obtain G .

A simple example illustrates: Suppose that $H = \sum_{i=1}^m h_i$ and P is an average price. By table 1, the h functions are of the general form

$$h_i(p_i, w) = \alpha_i(w) + \frac{\hat{\beta}(w)}{m} p_i \quad i = 1, \dots, m.$$

Differentiating the resulting (11) with respect to p_j for ($j \neq i$) implies

$$\frac{\partial^2 G}{\partial P^2} = 0,$$

or that the aggregate response function must be linear in the aggregate price. As long as H is linear G always inherits the same form as the h_i functions. As another example, we have already shown above for the Laspeyres price index and H linear that

$$G(P, w) = \frac{\alpha(w)}{\beta(w) - P},$$

which is to be compared with the associated h_i 's shown above.

Testing Aggregation Restrictions

As indicated above, choice of a supply aggregation rule and a price aggregator conjointly imply that micro supply equations are restricted to certain families of functional forms. For ex-

ample, table 1 suggests that, if aggregate supply equals the sum of firm supplies while the aggregate price represents the average price across firms, then firm-level supply equations must assume the Gorman polar form:

$$z_i(p_i, w) = \theta_i(w) + \beta(w)p_i.$$

So one way to test whether the sum of firm supplies and the average price aggregate consistently is to test whether the Gorman polar form accurately portrays the variation of the microeconomic price and supplies. The Gorman polar form is a special case of the general quadratic function,

$$y_i(p, w) = \theta_i(w) + \beta_i(w)p_i + \gamma_i(w)p_i^2.$$

Because the quadratic function provides a second-order flexible approximation of any differentiable $y_i(p, w)$, it has proved a popular representation (Shumway) especially when one is interested in testing for the linear special case. (The popular translog function, having the Cobb-Douglas as the linear special case, can also be interpreted as a quadratic function.) Using this approach, a statistical test for the validity of the linear aggregation hypothesis could be tested within the framework of a second-order flexible maintained model by testing the restrictions statistically,

$$\beta_i(w) = \beta(w) \quad i = 1, \dots, m$$

$$\gamma_i(w) = 0 \quad i = 1, \dots, m,$$

against the more general quadratic model.

Our general procedure for testing aggregation restrictions follows a similar approach: once a specific form is deduced as being required by the choice of the aggregate quantity and the aggregate price, nest that form within a more general form which is at least second-order flexible in prices, and then test the specific model against the more general form. To illustrate the procedure, we consider the conditions required to permit aggregation of wheat supplies and wheat prices across states using the aggregation rule that USDA uses to calculate its season average price for wheat as reported in *Agricultural Statistics*. The season average price reported in that publication is given by

$$P = \frac{\sum_{i=1}^{50} p_i y_i}{\sum_{i=1}^{50} y_i},$$

where y_i is now wheat supply in state i and

p_i is the estimated average price of wheat in state i .

A number of supply aggregates could be chosen to be combined with P , but one that is particularly attractive is the implicit supply aggregate (Y) defined by Fisher's product test:

$$Y \cdot P = \sum_{i=1}^{50} p_i y_i.$$

Of course, this quantity aggregate is just the sum of state supplies so that

$$Y = \sum_{i=1}^{50} y_i.$$

As noted above, Pope and Chambers have shown that the only family of supply functions consistent with this form is

$$y_i = \frac{\alpha_i}{\beta - p_i} \quad (i = 1, \dots, m).$$

To apply our procedure, we first rewrite this equation in the equivalent implicit form

$$(12) \quad y_i(\beta - p_i) = \alpha_i \quad (i = 1, \dots, m).$$

Because it has only two parameters, this equation is at most first-order flexible in prices. However, in a manner analogous to the discussion of the Gorman polar form, we note this form is a special case of the second-order flexible form:

$$(13) \quad y_i(\beta_i - p_i) + \gamma_i \cdot (\beta_i - p_i)^2 = \alpha_i \quad i = (1, \dots, m).$$

Because this form has three parameters (α_i , γ_i , β_i), it is potentially second-order differential flexible.

Within this maintained model, a test for the aggregation restrictions then becomes a test of the linear parametric restrictions,

$$(14) \quad \begin{aligned} \beta_i &= \beta \quad i = 1, \dots, m \\ \gamma_i &= 0 \quad i = 1, \dots, m. \end{aligned}$$

To execute this test, we employed a data set consisting of observations on wheat supply and wheat price for one state from each of the four main production regions in the United States for the period 1958–87. The states were Washington, Oklahoma, South Dakota, and Pennsylvania. We proceeded as follows: first, we fit the general form of the implicit supply equation to the data; second, we tested whether the form

$$(15) \quad y_i(\beta_i - p_i) = \alpha_i,$$

i.e., $\gamma_i = 0$, represented an acceptable explanation

of the price and quantity covariation in the model; and we then tested whether within this latter model the aggregable form (12) represented an acceptable restriction of the data. A direct test of the acceptability of this last version of the model against the most general version was also conducted.

The statistical model appended an additive error to the implicit supply equation represented above. The error term was assumed to have mean zero and be intertemporally independently distributed but with contemporaneous correlation across states. The estimation procedure was systems maximum likelihood and the results are reported in table 2. We first determine whether the aggregation restrictions represent a reasonable depiction of the sample data. When these seven restrictions are imposed upon the model ($\beta_i = \beta$, $\gamma_i = 0$ for all i), the log of the likelihood function decreased from -529.5148 to -553.1016 . Using an asymptotic chi-square test would lead to the rejection of the null hypothesis of the aggregation restrictions at all reasonable levels of confidence.

It is interesting to note, however, that the failure to accept the aggregation restrictions appears to emerge from the fact that the form

$$y_i(\beta_i - p_i) = \alpha_i$$

is a statistically inferior representation of the covariation in the data than the more general quadratic form. When this latter model is treated as the maintained model, one cannot reject the null hypothesis that $\beta = \beta_i (i = 1, \dots, m)$.

These statistical results suggest that

$$\begin{aligned} Y &= \sum_{i=1}^m y_i \text{ and} \\ P &= \frac{\sum_{i=1}^m y_i p_i}{\sum_{i=1}^m y_i} \end{aligned}$$

represent unacceptable supply aggregates for the current data. However, it is now possible to use the results contained in table 1 to construct alternative aggregates which are more consistent with the data. Suppose, for example, that we decide to retain $Y = \sum_i y_i$ as the supply aggregate. Table 1 and the results reported in table 2 then indicate that the aggregate price must belong to the general family

$$P = P^* \left(\sum_{i=1}^m \tau_i + \delta \left(\frac{\alpha_i}{\beta_i - p_i} - \gamma_i(\beta_i - p_i) \right); w \right).$$

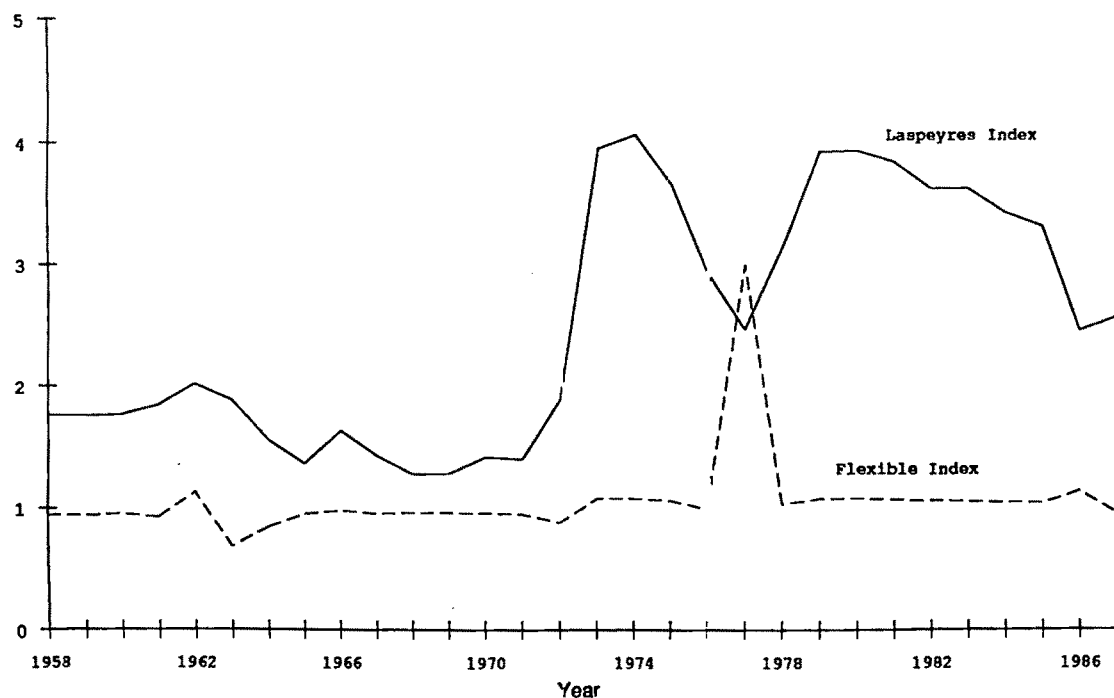
Table 2. Laspeyres Aggregation of Washington, Oklahoma, South Dakota, and Pennsylvania Wheat Production

Parameter	Unconstrained Estimates ^a	Constrained Estimates ^b	Constrained Estimates ^c
α_w	25.38 (28.51)	16.10 (29.53)	-32.89 (24.83)
α_o	-32.55 (26.41)	30.08 (33.48)	-26.86 (27.61)
α_s	7.95 (14.14)	4.59 (12.17)	-5.44 (10.82)
α_p	-4.74 (1.86)	-8.98 (2.60)	2.50 (2.21)
γ_w	30.85 (6.39)		
γ_o	59.45 (9.46)		
γ_s	10.23 (5.07)		
γ_p	1.72 (0.34)		
β_w	2.66 (0.13)	2.91 (0.19)	2.46 (0.09)
β_o	1.93 (0.16)	2.89 (0.19)	2.46 (0.09)
β_s	2.53 (0.14)	2.65 (0.14)	2.46 (0.09)
β_p	1.58 (0.12)	1.45 (0.19)	2.46 (0.09)
Log likelihood	-529.51	-542.48	-553.10

^a These are estimates of (13), where subscripts denote Washington, Oklahoma, South Dakota, and Pennsylvania, respectively. Standard errors are reported below each coefficient.

^b Constrained estimates are estimates of (15).

^c Estimates of (12); that is, all of the restrictions in (14) are imposed on (13).

**Figure 2. Two wheat price indexes**

For the sake of concreteness, take $\tau_i = 0$ ($i = 1, \dots, m$), $\delta = 1$, and P^* to be the exponential function. Figure 2 presents a graphical depiction of this aggregate price for the four sample states and compares it with the similarly computed Laspeyres index.

The preceding test has supposed that micro-economic data were available to test the aggregation hypothesis. Unfortunately, the empirical reality is that researchers frequently have only aggregate data with which to work. If that is the case, then the aggregation restrictions must be tested at the aggregate level. In an earlier section, we have shown that if the above price and quantity aggregates are used, then the aggregate supply relation must assume the same general form as the micro supply equations, i.e.,

$$Y = \frac{\alpha(w)}{\beta - \hat{P}}.$$

Therefore, a test for the aggregation restriction using aggregate data can be constructed along similar lines to those above, i.e., test this form (in implicit form) against the more general form,

$$(16) \quad Y(\beta - \hat{P}) + \gamma \cdot (\beta - \hat{P})^2 = \alpha(w).$$

Estimated versions of the aggregate model are presented in table 3. Estimates were obtained by the same procedure used for the micro forms. The aggregation restriction is that $\gamma = 0$. Because we now proceed as if no data are available on supply-price covariation at the microeconomic level, the aggregation test reduces to testing for the validity of the aggregate form versus (16). But because the asymptotic t -statistic on the estimate of γ exceeds 5, the hypothesis of aggregability is rejected at this level as well. Both macro and micro tests result in the same conclusion. The data are inconsistent with aggregation using the Laspeyres form.

Table 3. Laspeyres Aggregation—The Aggregate Equation

Parameter	Unconstrained Estimates ^a	Constrained Estimates ^b
α	-117.20 (56.44) 126.74	949.41 (181.56)
γ	(22.91) 1.15	5.74
β	(0.34)	(0.57)
Log likelihood	-181.32	-203.75

^a This column contains estimates of (16).

^b Column 3 contains estimates of (16) with γ constrained at zero.

Although this data set has led to the rejection of the Laspeyres aggregation hypothesis, it is clear that an acceptance of the aggregation hypothesis using aggregate data would have enabled us to identify a representation of the micro supplies consistent with the aggregate data. By the above, if aggregation had been accepted, then

$$\sum_{i=1}^m p_i y_i = -\alpha(w) + \beta(w) \sum_{i=1}^m y_i,$$

so that one representation of each $y_i(p_i, w)$ consistent with the data would be

$$y_i(p_i, w) = \frac{\alpha(w)/m}{\beta(w) - p_i}.$$

Conclusion

Aggregation across commodity groups has received considerable attention in agricultural economics. Such attention has focused in large part on testing for separable commodity groups in input, output, or price space and attempting to discover the nature of the aggregate index.

Similar in spirit (but not form) is the analysis of aggregate price indexes where aggregation is across economic entities. Because this form of aggregation is widespread, the practical implications of methods for more accurate measurement are large.

We have considered the implications of aggregation on aggregate response and the form of aggregate indexes. We have shown that exact aggregation imposes a definite structure on supplies, demands, or profits. In particular, commonly used price indexes imply specific families of responses with direct implications for elasticity measurement. These implications can be used to construct tests for exact aggregation.

By way of an empirical inquiry on wheat supply, the most commonly used and reported index of aggregate price was not supported by exact aggregation. Based upon these findings, we proposed, calculated, and reported one of many possible empirically based indexes which are consistent with the supply-price covariation in the data. The testing procedures presented here can be used directly to aid in the discovery of simple but empirically relevant aggregates for all agricultural price series.

[Received February 1989; final revision received September 1990.]

References

- Antle, J. M. "Econometric Estimation of Producers' Risk Attitudes." *Amer. J. Agr. Econ.* 69(1987):509-22.
- Ball, V. E. "Modeling Supply Response in a Multiproduct Framework." *Amer. J. Agr. Econ.* 70(1988):813-25.
- Blanciforti, L., and R. Green. "An Almost Ideal Demand System Incorporating Habits: An Analysis of Expenditures on Food and Aggregate Commodity Groups." *Rev. Econ. and Statist.* 65(1983):511-15.
- Chambers, R. J. *Applied Production Analysis*. Cambridge: Cambridge University Press, 1988.
- Chavas, J.-P., and C. Cox. "A Nonparametric Analysis of Agricultural Technology." *Amer. J. Agr. Econ.* 70(1988):303-10.
- Courant, R. *Differential and Integral Calculus, Vol. II*, rev. ed. New York: Interscience Publishers, 1937.
- Deaton, A., and J. Muellbauer. *Economics and Consumer Behavior*. Cambridge: Cambridge University Press, 1980.
- Eales, J., and L. Unnevehr. "Beef and Chicken Product Demand." *Amer. J. Agr. Econ.* 70(1988):521-32.
- Gorman, W. M. "Community Preference Fields." *Econometrica* 21(1953):63-80.
- . "Compatible Indices." *Econ. J.* 96, supplement (1968), pp. 83-95.
- Gould, B., and W. E. Saupe. "Off-Farm Labor Market Entry and Exit." *Amer. J. Agr. Econ.* 71(1989):960-69.
- Green, R., and J. Alston. "Elasticities in AIDS Models." *Amer. J. Agr. Econ.* 72(1990):442-45.
- Jorgenson, D., L. J. Lau, and T. M. Stoker. "Aggregate Consume: Behavior and Individual Welfare." *Macroeconomic Analysis: Essays in Macroeconomics and Econometrics*, ed. D. Currie, R. Novay, and D. Peel. London: Croom Helm, 1981.
- Muellbauer, J. "Aggregation, Income Distribution, and Consumer Demand." *Rev. Econ. Stud.* 62(1975):525-43.
- Nichols, J. N. "Testing a Theory of Exact Aggregation." *J. Bus. and Econ. Statist.* 7(1989):259-65.
- Pope, R. D., and R. G. Chambers. "Price Aggregation Over Price-Taking Firms." *Rev. Econ. Stud.* 56(1989):297-309.
- Robison, L. J., D. A. Lins, and R. VenKataraman. "Cash Rents and Land Values in U.S. Agriculture." *Amer. J. Agr. Econ.* 67(1985):794-805.
- Shumway, C. R. "Supply, Demand, and Technology in a Multiproduct Industry: Texas Field Crops." *Amer. J. Agr. Econ.* 65(1988):748-60.
- Stoker, T. M. "Completeness, Distribution Restrictions, and the Form of Aggregate Functions." *Econometrica* 52(1984):387-907.
- U.S. Department of Agriculture. *Agricultural Statistics*. Washington DC, various years.
- Weaver, R. L. "Multiple Input, Multiple Output Production Choice, and Technology in the U.S. Wheat Region." *Amer. J. Agr. Econ.* 65(1983):45-56.

On Nonlinear Dynamics: The Case of the Pork Cycle

Jean-Paul Chavas and Matthew T. Holt

New methods for analyzing nonlinear dynamic processes are used to evaluate the hog-corn price ratio. The results present evidence of nonlinear dynamics in the pork cycle. Moreover, while GARCH processes account for some of the nonlinearities, the pork cycle is apparently characterized by more complex dynamic forms. The empirical analysis provides some evidence of the presence of chaos in the pork cycle. The results also indicate that the dynamic process generating the pork cycle is nonlinear and cannot be adequately characterized by a small number of state variables.

Key words: business cycle, chaos, GARCH, nonlinear dynamics.

The pork market has been viewed by economists as a clear example of a business cycle. Beginning with Coase and Fowler, an extensive body of research has focused on identifying the source and nature of the pork cycle (e.g., Breimyer, Larson, Jelavich, Jameson). For example, Harlow proposed to explain a four-year hog cycle in the context of the cobweb theorem and a delayed supply response to market price. Using harmonic analysis, Talpaz argued that hog prices follow several different cycles, each with a different period and magnitude. Using linear time-series models, Wallis, Shonkwiler and Spreen, Bessler and Kling, and Kaylen presented empirical evidence on the existence and nature of the pork cycle. Although the results often vary according to the period analyzed and/or the methodology used, the existence of a pork cycle has been strongly supported by the empirical evidence.

In many respects, the continuing presence of any price cycle is disturbing for economists. If a predictable price cycle exists, then producers responding in a countercyclical fashion could earn larger than "normal" profits over time (Hayes and Schmitz). Such profits could occur even with lags in the production process because predictable price movements would still influence pro-

duction decisions. Eventually, countercyclical production response would smooth out price fluctuations at the market level, causing the cycle to disappear.

The implication is that predictable price cycles are incompatible with rational producer behavior. Hence, the continuance of a business cycle could imply that producers are not behaving rationally and are ignoring important information. The assumption of naive producer behavior, while dating back to Ezekiel and others, still is a basic premise in much of the literature on livestock supply response.

An alternative explanation for the existence of a business cycle is that the cycle itself is not perfectly predictable. The law of motion underlying the business cycle may be a deterministic nonlinear relationship that generates unpredictable patterns (e.g., Grandmont and Malgrange). In recent years, examples of nonlinear, deterministic dynamic systems that produce seemingly random behavior have been reported in both the physical and social sciences (May, Grandmont and Malgrange). Such systems, usually called "chaotic," are necessarily associated with nonlinear dynamics (Hao, Baumol and Benhabib).

Until recently, economists have not considered that unpredictable patterns could result from deterministic, nonlinear dynamics. Recent work by Day; Brock; Scheinkman and LeBaron; Brock and Sayers; Frank, Genzay, and Stengos and others has, however, begun to examine the prospect that observed economic events could result from deterministic chaos. This research

Jean-Paul Chavas and Matthew T. Holt are a professor and an assistant professor, Department of Agricultural Economics, University of Wisconsin.

Financial support was provided by a Hatch grant from the College of Agriculture and Life Sciences, University of Wisconsin.

The authors would like to thank two anonymous reviewers for helpful comments on an earlier draft of this paper.

has resulted in a set of tools and statistical tests useful for investigating deterministic nonlinear dynamics in economic analysis.

The objective of this paper is to consider whether a particular business cycle, the pork cycle, may be characterized by a deterministic nonlinear dynamic process, and hence may not be fully predictable. In so doing, we present some of the analytical tools currently available for nonlinear dynamic analysis and illustrate their application to the pork cycle. In the next section deterministic nonlinear dynamic systems are briefly reviewed. Methods for detecting nonlinear dynamics in economic data are discussed; and the dynamics of the hog-corn price ratio in the United States is investigated.

Nonlinear Dynamics

The focal point of most dynamic analysis is a system of difference equations which dictate how a vector of state variables, y_t , evolves over time, t .¹ In general form, a system of n first-order difference equations can be written as

$$(1) \quad y_{t+1} = f(y_t),$$

where y_t is an n -vector of state variables and the function f consists of n separate linear or nonlinear component functions. The system of first-order difference equations in (1) provides a basis for analyzing any discrete-time dynamic system; any system of second- or higher-order difference equations can always be reformulated as an equivalent first-order system (Luenberger, p. 96).

Linear State Equations

The properties of the state equations (1) are best understood in the context of linear models where $f(y_t) = Ay_t$, and A is an $(n \times n)$ matrix. In this situation, an analytical solution to the time path of the state variables can be obtained (Chow). It is well known that linear models result in state trajectories that can exhibit several types of behavior: trajectories that converge to a steady state, trajectories that become arbitrarily large over time, or periodic trajectories that oscillate or cycle.

For linear systems, the type of behavior can be deduced from the characteristic roots of the system matrix, A . Complex roots imply oscillations and cyclical patterns (e.g., Luenberger, p. 158). The system will also be stable (unstable) whenever the modulus of the dominant characteristic root is less than (greater than) one.

Because linear models exhibit well-known analytical properties and are empirically tractable, they have been used extensively to investigate dynamic behavior in economics. Indeed, time-series analysis and the analysis of distributed lags in econometrics have mostly used linear models. Also, most studies of pork market dynamics have used linear models (e.g., Wallis, Shonkwiler and Spreen, Bessler and Kling, Kaylen).

Linear models can provide useful insights into the dynamic properties of economic systems. But, are economic models always best characterized with linear state equations? Many economists would agree that dynamic relationships in economics are often fundamentally nonlinear, but that linear systems provide a useful "first-order" approximation to the true (but unknown) laws of motion. This argument is consistent with the methods frequently used to analyze the stability of nonlinear systems. That is, according to Lyapunov's first method, any system of nonlinear differentiable state equations can be linearized at a point (Luenberger, p. 324). The analytical results for inferring "global" stability in linear systems can then be used to infer "local" stability in nonlinear systems.² Several implications of using linear systems to approximate nonlinear systems are discussed below.

Nonlinear State Equations

By comparison, relatively little research has focused on nonlinear dynamics in economics. Unlike linear models, tractable methods for solving and analyzing nonlinear systems are not generally available. When analyzing nonlinear models, it is often necessary to obtain results numerically. Consequently, recent progress in computer technology has created new opportunities for investigating the dynamic properties of nonlinear state equations.

Quantitative predictions can differ between

¹ The discussion is limited to discrete-time dynamic systems. Similar results could be obtained for continuous-time systems. In this case, the dynamic equations (1) would take the form of first-order differential equations instead of difference equations.

² The term local is used when evaluating the stability of nonlinear systems because the results hold only in some neighborhood of the point of approximation.

systems of linear and nonlinear state equations. "Approximation errors" are introduced if a linearized model is used to represent a nonlinear system. Such considerations arise in policy analysis or in forecasting when dynamic simulations are used to predict values for the state variables that are well beyond the observed range of the data.

Moreover, nonlinear state equations can exhibit dynamic behavior that is qualitatively different from that of linear state equations. Nonlinear state equations can exhibit dynamic patterns that cannot, under any circumstances, be generated or reproduced by a system of linear state equations. These new and different patterns have been grouped in the literature under the general term "chaos."

In the early 1960s, Lorenz first identified chaotic patterns in the context of a dynamic weather model. He discovered that a deterministic system of three differential equations could generate nonperiodic but bounded solutions. In 1975, Li and Yorke introduced the term "chaos" to denote the apparently random output of certain nonlinear dynamic systems. In 1976, May cited the complicated dynamics generated by simple population models. Over the last fifteen years, there has been growing interest across disciplines (e.g., physics, climatology, ecology, economics) in nonlinear deterministic models and the highly irregular trajectories they can generate.

The word "chaos" refers to "unpredictability" in completely deterministic dynamic systems. The concept is at odds with the more traditional view that accurate predictions are always possible from a deterministic model. Chaos means, for example, that apparently random fluctuations in some state variables are not necessarily associated with a stochastic process or with errors in measuring the model's parameters (e.g., Brock). The potential for chaos warns us of the dangers of extrapolation and of the difficulties of economic forecasting. It also suggests a need for a careful reevaluation of the traditional way of thinking about stochasticity in modeling. Finally, chaos should not be equated with disorder. Rather, chaos is considered a kind of order without periodicity.

Unfortunately, few analytical results are available for characterizing the potential for, and nature of, chaos associated with a general system of state equations as in (1). Most analytical results have been obtained using particular specifications of (1) (e.g., Hao). Apparently, nonlinearities can appear in so many ways that it is

difficult to develop a general criterion or classification scheme for characterizing chaos. Nonetheless, chaos occurs in a variety of nonlinear dynamic models that are often relevant in explaining a fairly wide class of observable events (Hao).

Deterministic Chaos, an Example

The above discussion suggests that chaos is best illustrated in the context of specific models. One useful example is given by the deterministic state equation

$$(2) \quad y_{t+1} = \alpha y_t(1 - y_t),$$

where y_t is a scalar and α is a positive "tuning" parameter. Equation (2) is the logistic difference equation discussed by May. It is one of the simplest nonlinear state equations possible: it is a first-order quadratic difference equation with a single state variable. However, as shown by May, this simple deterministic state equation can generate complex dynamics. The time path of y_t in (2) is determined completely by the parameter α and by the initial condition, y_0 . The logistic equation (2) is unstable for any $y_0 < 0$ or $y_0 > 1$. We focus on the case where $0 < y_t < 1$ and $0 < \alpha < 4$. In these intervals, the dynamic properties of (2) depend entirely on the value of the tuning parameter, α .

For $0 < \alpha < 3$, model (2) is stable and converges to a unique steady state for any initial condition y_0 , $0 < y_0 < 1$. For $3 < \alpha < 3.5699$, the logistic equation exhibits cyclical patterns. For any initial condition, $0 < y_0 < 1$, the state variable converges on stable cycles of period 2^n , where n is an integer that depends on α . For example, if $3 < \alpha < 3.4495$, then $n = 1$ and the model would exhibit a two-period cycle. As α increases from 3.4495 to 3.5699, n takes on successively the values of 2, 3, 4, ..., each value corresponding to a stable cycle of period 4, 8, 16, ... (see May). In other words, as α increases from 3 to 3.5699, first there is one stable two-period cycle; then a four-period cycle emerges; then an eight-period cycle appears, and so on, ad infinitum. Given $0 < \alpha < 3.5699$, model (2), therefore, exhibits either stable or cyclical patterns in the interval $0 < y_t < 1$. That is, the state variable y_t would eventually converge to its "attractor": the point or set of points (e.g., limit cycle) describing the stable equilibrium of the system. Such predictable patterns can be found in linear models and are not specific to nonlinear models.

The dynamic properties of (2) are of particular interest, however, when $3.5699 < \alpha < 4$. In this region, model (2) exhibits "chaotic" behavior. Chaos implies that "small" changes in the initial condition, y_0 , can give rise to time paths that eventually diverge. Such divergence has been called "sensitive dependence to initial conditions" and is illustrated numerically by Baumol and Benhabib (pp. 93–95) for the logistic equation (2). Sensitivity to initial conditions shows that long-term predictions from a chaotic system are virtually impossible. This is the point first made by Lorenz, who argued that weather patterns are deterministic but chaotic, implying the impossibility of accurate long-term weather predictions.³

The divergence of neighboring paths does not preclude the convergence of the state variables to a bounded set of points. This bounded set of points is called a "chaotic attractor." More specifically, a chaotic attractor is an uncountable set of points attracting all paths starting within the neighborhood of the set, where such paths are aperiodic. Aperiodicity here means the existence of an infinite number of different periodic cycles that never repeat themselves. Thus, under chaos, an uncountable number of initial points y_0 gives totally aperiodic (although bounded) trajectories.

One implication of nonlinear dynamics is that a chaotic deterministic system can generate a time series that appears completely random. This, in turn, raises the issue of distinguishing between a random process and a chaotic deterministic process.⁴ For example, standard statistical tests relying on the spectrum or on the autocovariance function may suggest that a chaotic trajectory is indistinguishable from a trajectory generated by a "white noise" stochastic process (Brock and Sayers). In this case, standard tests for "randomness" could not discriminate between a stochastic explanation and a deterministically chaotic explanation of a time series. This indicates an important weakness of the traditional linear statistical approach to time-series

analysis. It also suggests a need for different methods in the analysis of nonlinear dynamic processes.

Tools for Nonlinear Dynamic Analysis

Nonlinear dynamic analysis is fairly new; it has received serious attention in the scientific community only during the last fifteen years. Consequently, the analytical tools for investigating nonlinear dynamic systems are still being developed. In this section, we briefly review some of the tools currently available and discuss their potential for empirical analysis.

Consider a univariate time-series of length T , represented by the scalar x_t . We wish to investigate the nonlinear dynamics of x_t , where

$$(3a) \quad x_t = h(y_t),$$

$$(3b) \quad y_{t+1} = f(y_t), \quad t = 1, 2, \dots, T.$$

Here, (3a) is the observer equation, (3b) is the equation of motion, and y_t is an $(n \times 1)$ vector of state variables characterizing the dynamics of x_t . Assume that the x_t 's are observed but the state variables y_t are not. The problem then is how to use observations on the x_t 's to investigate the underlying dynamic structure of the state variables. We assume that the functions $h(\cdot)$ and $f(\cdot)$ in (3) are differentiable and that the process $y_{t+1} = f(y_t)$ possesses a unique attractor. Three questions related to the equation of motion (3b) are of interest: (a) how can one tell whether the dynamic properties of x_t are chaotic? (b) how can one uncover the dimension, n , of the state vector y_t that underlies the dynamic properties of x_t ? And (c) how can one test for the existence of nonlinearities? Each issue is addressed in turn.

Identifying Chaos

Chaos is associated with the local divergence of neighboring paths in the attractor. Given this local property, consider the linearized version of model (1) around some resting point, \bar{y} :

$$y_{t+1} = \bar{y} - A(\bar{y})\bar{y} + A(\bar{y})y_t,$$

where $A(\bar{y}) = \partial f / \partial y(\bar{y})$. In principle, the modulus of the characteristic roots of $A(\bar{y})$ could be averaged over a trajectory to obtain a set of global measures.⁵ Such measures are called "Lyapu-

³ Using his simple three-equation nonlinear deterministic model, Lorenz discovered that long-term simulations were extremely sensitive to the initial conditions, thus implying the unpredictability of long-term weather patterns.

⁴ To the extent that random variables are used to represent part of the real world we do not understand well, the issue may not be to choose between a random versus a deterministic explanation of observable phenomena. Rather, the motivation should be to obtain better explanations for the seemingly "unpredictable" events that often occur in the real world. Chaotic models may help improve understanding of certain phenomena beyond that which is possible for stochastic models.

⁵ The characteristic roots of $A(\bar{y})$ determine the local stability of the state equation (1). The system is locally stable (unstable) if the modulus of the dominant root of $A(\bar{y})$ is less than (greater than) one.

nov exponents": they are generalizations of the modulus of characteristic roots along general forward trajectories. In the case of a single state system, $n = 1$, $A(\bar{y})$ is a scalar, and there is only one Lyapunov exponent defined as

$$\lambda = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{i=1}^T \ln(|A(y_i)|).$$

From the above discussion, $\lambda > 0$ corresponds to chaotic trajectories ($|A(\cdot)| > 1$ identifying local divergence) while $\lambda < 0$ corresponds to nonchaotic trajectories ($|A(\cdot)| < 1$ identifying local stability). In the general case of n state variables, there exist n Lyapunov exponents, $\lambda_1, \lambda_2, \dots, \lambda_n$.⁶ Supposing that $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$, then λ_1 denotes the largest Lyapunov exponent. For inferring chaos, it follows then that a positive largest exponent, $\lambda_1 > 0$, corresponds to chaotic trajectories, while $\lambda_1 < 0$ corresponds to nonchaotic trajectories. Thus, an empirical estimate of the largest Lyapunov exponent λ_1 will provide a measure of the existence of chaos in dynamic analysis.

Wolf et al. proposed a numerical algorithm for calculating the Lyapunov exponents of a time series. While the statistical properties of their estimator are not yet known, the Wolf et al. es-

is given in part by Takens' theorem, which states that the dynamics of m -histories x_t^m is equivalent to the dynamics of y over its attractor when $m \geq 2n + 1$ (Takens). This indicates that the smallest unit useful for dynamic analysis is given by x_t^m in (4) with $m \geq 2n + 1$.

Second, if the unobserved state equation (1) is identifiable, one way to measure the "dimension" of the attractor of y_t is to use the generalized definition of dimension proposed by Hausdorff. (See Grassberger and Procaccia for details.) In general, the Hausdorff dimension D of the attractor of y_t is at most equal to n , the dimension of the state vector y_t . However, the Hausdorff dimension D of a chaotic attractor could be a non-integer and smaller than n . An often cited example of a non-integer dimension (also called fractal dimension) is the Cantor set which has dimension $D = .6309$.⁷

Although the Hausdorff dimension D has received much attention in the literature, it is difficult to compute empirically. Consequently, Grassberger and Procaccia (GP) proposed the "GP correlation dimension" α as an alternative measure

$$(5) \quad \alpha = \lim_{\epsilon \rightarrow 0} \ln(C_m(\epsilon)) / \ln(\epsilon),$$

where

$$(6) \quad C_m(\epsilon) = \lim_{T \rightarrow \infty} \{\text{number of pairs } (i, j) \text{ whose distance, } \|x_i^m - x_j^m\|, \text{ is less than } \epsilon, i \neq j\} / (T^2 - T).$$

timator of λ_1 does provide a basis for detecting the presence of chaos in a given time series.

The Dimension of the Attractor

Noting from (3) that $y_{t+2} = f(f(y_t)) = f^2(y_t)$ and $y_{t+k} = f(f^{k-1}(y_t)) = f^k(y_t)$, $k = 1, 2, \dots$, consider the m -history

$$(4) \quad \begin{aligned} x_t^m &= (x_t, x_{t+1}, \dots, x_{t+m-1}) \\ &= (h(y_t), h(f(y_t)), \dots, h(f^{m-1}(y_t))) \\ &= J_m(y_t), \end{aligned}$$

where $J_m(y_t)$ is a mapping from R^n to R^m . Because x_t is observed but y_t is not, the first question addresses the minimum number of observations on x_t that are needed to recover the unobserved state equations in (3). The answer

Grassberger and Procaccia show that the GP correlation dimension α is a lower bound estimate of the Hausdorff dimension, where $\alpha \leq D$. They also argue that, in most cases, $\alpha \cong D$, implying that α can provide a useful measure of the underlying structure of the attractor of y_t . The empirical estimation of the GP correlation dimension is also computationally convenient and has several nice properties (Brock). Even though the statistical properties of this estimator are unknown at this time, the GP measure α is a useful method for identifying the minimum number of state variables needed in the dynamic analysis of a time series.

⁶ See Benettin, Galgani, and Strelcyn for a general definition of Lyapunov exponents and for a proof of their existence.

⁷ To see this, consider a Cantor set on the interval (0, 1) obtained by deleting the central 1/3, i.e., the interval (1/3, 2/3). Then repeat this operation with respect to the remaining segments an infinite number of times. In this case, increasing the linear size of the interval (0, 1/3) by a factor of 3 yields 2 copies of the same object. The resulting dimension is $D = \ln 2 / \ln 3 = .6309$. This indicates that the volume of a Cantor set increases by 63.09% given a doubling of each of its spatial dimensions.

A Whiteness Diagnostic Test

As discussed in the second section, traditional statistical tests may not distinguish between a random process and a deterministic chaotic process. However, Brock proposed a statistical test (the BDS test) to empirically analyze this distinction. The BDS test statistic is

$$(7) \quad W = T^{1/2}[C_m - (C_1)^m]/b_m,$$

where C_1 and C_m are given in (6) and b_m is the asymptotic variance of $[C_m - (C_1)^m]T^{1/2}$ (see Brock). Under the null hypothesis that x_t is an independent identically distributed (white noise) random variable, as $T \rightarrow \infty$, the W statistic in (7) converges to a standard normal variable with mean zero and variance one. Thus, the BDS test provides a statistical foundation for obtaining evidence against the null hypothesis that x_t is a white noise random process. Given the inability of more traditional statistical tests to distinguish between white noise and a deterministic but chaotic process, the BDS test can help uncover evidence of nonlinear dynamics. In particular, it can help detect misspecifications from using a linear model fit to nonlinear data.

The Residual Test Theorem

The results just discussed can be applied to any observable time-series x_t satisfying some weak regularity conditions. In applied work, the time series x_t is often assumed to follow a linear process that can be represented by an $AR(q)$ model of the form:

$$(8) \quad \beta_0 + B(L)x_t = e_t,$$

where $B(L) = \beta_1 L + \dots + \beta_q L^q$, L is a lag operator such that $L^s x_t = x_{t-s}$, β_0 is an intercept, and e_t is an innovation with mean zero and finite variance, h_t . Of particular interest here is Brock's theorem on the dynamic properties of the residual e_t in (8). Brock's residual test theorem states that, given a set of regularity conditions, the GP correlation dimension and the largest Lyapunov exponent of x_t and e_t are the same. This result suggests that the tools for analyzing nonlinear dynamics can be applied to the original data x_t , as well as to the innovations e_t associated with standard time-series models. This attractive characteristic increases the range of applications of the results discussed earlier. An empirical illustration of these tests and procedures is presented next.

An Application to the Pork Cycle

Here we investigate the nature of the U.S. pork cycle by analyzing the dynamics of the hog-corn price ratio. Quarterly observations on the U.S. hog-corn price ratio were obtained from various issues of *Livestock and Meat Statistics* for the period 1910–84. We first consider a typical linear time-series model where x_t , the hog-corn price ratio at time t , is assumed to follow an $AR(q)$ process as in (8), where e_t is normally distributed with mean zero and variance h_t . The order q of the AR process (8) was evaluated using Akaike's information criterion.

Equation (8) was estimated by the maximum likelihood method under two specifications: first, assuming that e_t is a homoscedastic white noise process where $h_t = h$, a scalar; and second, assuming that e_t follows a GARCH(1, 1) process where $h_t = a_0 + a_1 e_{t-1}^2 + a_2 h_{t-1}$ (Bollerslev).⁸ The Akaike criterion was minimized at $q = 20$. Table 1 presents the parameter estimates for an $AR(20)$ (where $h_t = h$), and for an $AR(20)$ with a GARCH(1, 1) error process. The estimates of $B(L)$ are similar in both instances. However, using a likelihood ratio test, the null hypothesis that $a_1 = a_2 = 0$ is rejected at the 5% level. Thus, statistical evidence indicates that the error term e_t is not homoscedastic and that its variance h_t is changing over time.

Following Luenberger (p. 96), the expected value of (3) can be expressed as an equivalent system of first-order difference equations: $X_t = HX_{t-1}$. The characteristic roots of the companion matrix, H , provide useful information about the dynamics of the estimated model (8). Details about the characteristic roots for both the estimated $AR(20)$ and the $AR(20)$ -GARCH(1, 1) models are presented in table 2. The results in table 2 indicate that both estimated models are dynamically stable and have cycles with periods ranging from one to five years. These results are consistent with previous research on cycles in the pork market (e.g., Wallis, Bessler and Kling Kaylen). Moreover, the presence of multiple cycles is also consistent with the results obtained by Talpaz.

To check the validity of the estimated versions of model (8), several standard tests were performed on the estimated residuals, e_t . The Box-Pierce test examines the estimated resid-

⁸ GARCH stands for "generalized autoregressive conditional heteroscedasticity."

Table 1. Parameter Estimates for the AR(20) and AR(20)-GARCH(1, 1) Models

	AR(20)		AR(20)-GARCH(1, 1)	
	Coefficient	Standard Error	Coefficient	Standard Error
Intercept	.5909	.0528	.4874	.5439
β_1	1.1046***	.0613	1.1032***	.0667
β_2	-.3542***	.0915	-.3446***	.0969
β_3	.0501	.0936	.0806	.0966
β_4	-.0152	.0931	-.0844	.0920
β_5	-.2339***	.0931	-.1942**	.0898
β_6	.1506	.0938	.1437*	.0915
β_7	.0629	.0942	.0562	.0934
β_8	-.0083	.0941	.0246	.0936
β_9	-.1804**	.0933	-.2105**	.0922
β_{10}	.1573*	.0940	.1500*	.0960
β_{11}	-.0639	.0940	-.0437	.0910
β_{12}	.1959**	.0936	.1734**	.0940
β_{13}	-.1131	.0944	-.0824	.0938
β_{14}	-.0549	.0948	-.0691	.0918
β_{15}	.1898**	.0945	.1829**	.0916
β_{16}	.0568	.0944	.0603	.0903
β_{17}	-.1850**	.0943	-.1757**	.0893
β_{18}	.2040**	.0951	.2038**	.0907
β_{19}	-.2021**	.0938	-.1725**	.0879
β_{20}	.2043***	.0642	.1714***	.0612
GARCH, a_0	2.5088***	.0555	.4016	.3156
GARCH, a_1			.1065*	.0709
GARCH, a_2			.7378***	.1652
Log likelihood	-264.482		-260.711	
	$R^2 = .844$		$R^2 = .928^a$	

Note: Asterisks indicate significance levels: *** = 1%, ** = 5%, * = 10%.
^a The R^2 for the AR(20)-GARCH(1, 1) model was computed using the standardized residuals.

uals for white noise.⁹ The white noise test proposed by Ljung and Box is a modification of the first test.¹⁰ Last, McLeod and Li proposed a test for white noise which is based on the squared-residual autocorrelations of ϵ_t .¹¹ This test is use-

⁹ The Box-Pierce Q_a and the Ljung-Box Q_a^* tests are based on the residual autocorrelation function:

$$r_a(k) = \frac{\sum_{t=k+1}^T \epsilon_t \epsilon_{t-k}}{\sum_{t=1}^T \epsilon_t^2},$$

where the ϵ_t are the innovations associated with equation (8). Under the null hypothesis that the residuals are white-noise, the Box-Pierce statistic

$$Q_a = T \sum_{i=1}^m r_a^2(i)$$

is distributed approximately as a $\chi^2(m - q)$.
¹⁰ Ljung and Box proposed to use the statistic

$$Q_a^* = T(T + 2) \sum_{i=1}^m r_a^2(i) / (T - i),$$

which is approximately distributed as a $\chi^2(m - q)$ under the null-hypothesis of white noise.

¹¹ The squared-residual autocorrelation function is given by

$$r_{aa}(k) = \frac{\sum_{t=k+1}^T (\epsilon_t^2 - \sigma^2)(\epsilon_{t-k}^2 - \sigma^2)}{\sum_{t=1}^T (\epsilon_t^2 - \sigma^2)^2},$$

where

Table 2. Dynamics Associated with the Characteristic Roots of H

AR(20)		AR(20)-GARCH(1, 1)	
Modulus	Period ^a	Modulus	Period
.992	^b	.970	3.00
.966	^b	.960	^b
.951	4.51	.950	^b
.941	2.29	.949	1.82
.935	1.88	.945	2.27
.931	2.95	.929	1.44
.924	1.01	.924	5.14
.921	1.43	.917	1.05
.916	1.38	.900	4.67
.911	4.55	.885	1.34
.831	1.12	.794	1.05

Note: The roots are ordered according to their modulus.
^a The period is denoted in years.
^b These are real roots and hence do not exhibit cyclical behavior.

$$\sigma^2 = \sum_{t=1}^T \epsilon_t^2 / T.$$

McLeod and Li show that, under white noise, the statistic

$$Q_{aa} = T(T + 2) \sum_{i=1}^m r_{aa}^2(i) / (T - i)$$

is asymptotically distributed as a $\chi^2(m)$.

ful for detecting certain types of nonlinear statistical dependence in time-series residuals.

The above tests were used to determine whether the AR(20) residuals in (8) are white noise for $m = 40$. The test statistics were 23.34 and 26.49 for the Box-Pierce and Ljung-Box tests, respectively. Given a critical value of 31.41 from the asymptotic chi-square distribution at the 5% level, the null hypothesis that the residuals of (8) are white noise cannot be rejected in either case. Alternatively, the test statistic for the McLeod-Li test was 40.06. Given a critical value of 55.76 from the asymptotic chi-square distribution at the 5% level, the hypothesis that second-order autocorrelation is present in the residuals, e_t , also is rejected. Similar results were obtained for the standardized residuals from the AR(20)-GARCH(1, 1) model.¹² These tests suggest that the linear model (8) (with either homoscedastic errors or with GARCH errors) appropriately represents price dynamics in the hog market.

Now consider the BDS test. Choosing $m = 40$, the W statistics (7) were calculated for the original data x_t , for the AR(20) residuals in (8), and for the standardized residuals of the AR(20)-GARCH(1, 1) model. The results are reported in table 3 for selected values of ϵ . Recall that under the null hypothesis of white noise, W converges asymptotically to $N(0, 1)$. Thus, the critical value of the W test at the 5% level is 1.96.

The results in table 3 indicate that the null hypothesis that x_t is white noise is strongly rejected. The W statistics also provide statistical evidence that the AR(20) residuals are not, in fact, white noise. Given the results of the more

traditional tests for white noise, it follows that obtaining significant W statistics for the residuals of (8) provides important evidence of model misspecification caused by fitting a linear model to data generated by a nonlinear process. Moreover, these results are consistent with those of Brock and Sayers in their analysis of business cycles.

The BDS test applied to the standardized residuals from the AR(20)-GARCH(1, 1) model is also of interest. First, the evidence against white noise is not as strong as for the AR(20) residuals (see table 3). This result indicates that the GARCH specification is perhaps capturing some of the nonlinearities of the dynamic process. Second, for high values of ϵ , the BDS test still indicates that the standardized AR(20)-GARCH residuals are not white noise. Thus, the existence of nonlinear dynamics in the pork cycle cannot be fully explained by the GARCH error process. Importantly, this evidence for nonlinearities indicates that fundamental asymmetries may exist between the expansion phase and the contraction phase of the hog cycle.

Given that the hog-corn price ratio is associated with nonlinear dynamics, evidence on the nature of the underlying dynamic process also can be obtained by measuring the GP correlation dimension and the largest Lyapunov exponent. Brock and Sayers proposed using

$$(9) \quad SC_m(\epsilon_i) = (\ln C_m(\epsilon_i) - \ln C_m(\epsilon_{i-1})) / (\ln(\epsilon_i) - \ln(\epsilon_{i-1}))$$

as an empirical estimate of the GP correlation dimension, where C_m is defined in (6). The dimension estimates, SC_m , are reported in table 3 (using $m = 40$) for the hog-corn price ratio x_t ,

¹² The standardized residuals for the AR(20)-GARCH(1, 1) model are given by $e_t h_t^{-1/2}$, where h_t is the conditional variance.

Table 3. Dimension Estimates (SC_m) and BDS White Noise Test Statistics (W)

ϵ	x_t		e_t from AR(20)		Standardized e_t from AR(20)-GARCH(1, 1)	
	SC_m	W	SC_m	W	SC_m	W
.9 ²	.009	51.50	.431	8.329	.126	2.305
.9 ³	.218	41.85	1.181	5.300	.341	6.070
.9 ⁴	.724	25.90	1.924	2.463	.996	1.659
.9 ⁵	1.115	17.41	3.557	1.398	1.802	.215
.9 ⁶	1.391	16.03	4.077	.851	3.023	-.298
.9 ⁷	1.615	18.27	6.210	1.604	4.553	-.262
.9 ⁸	2.330	24.85	8.501	2.040	6.365	-.387
.9 ⁹	3.948	42.21	11.058	2.227	9.095	-.279
.9 ¹⁰	4.977	78.47	14.756	2.081	12.921	-.427
.9 ¹¹	5.591	184.59	25.603	2.035	21.582	-.926
.9 ¹²	9.115	554.04			25.911	-.909
.9 ¹³	9.339	161.970				

Note: All results are obtained for $m = 40$.

for the AR(20) residuals in (8), and for the standardized residuals of the AR(20)-GARCH(1, 1) model. The dimension estimates for x_t are rough: they can be as high as 9. The dimension estimates for both e_t from the AR(20) model and $e_t h_t^{-1/2}$ for the AR(20)-GARCH(1, 1) model are also rough but tend to be higher: they can be as high as 25. Based on the GP dimension estimates, the underlying dynamic process appears complex and would probably not be adequately represented by just a few state variables.

Last, given the evidence of nonlinearity provided by the BDS test, is there any evidence of chaotic behavior in the pork cycle? This question is investigated by calculating the largest Lyapunov exponent, λ_1 , using Wolf et al.'s algorithm.¹³ The estimates of λ_1 were positive: .0041 for x_t and .0038 for the AR(20) residuals. Thus, the estimates of λ_1 are in the chaotic region. Their small magnitude indicates that the rate of divergence of neighboring paths is low. Thus, the pork cycle may still be predictable in the short- and intermediate-run, even if chaotic. Only in the longer run may neighboring paths diverge and, hence, become unpredictable. Finally, the estimates of λ_1 are close to the non-chaotic region ($\lambda_1 < 0$). Because the statistical properties for the estimator of λ_1 are currently unknown, we cannot conclude there is strong evidence of chaos in the dynamics of the hog-corn price ratio.¹⁴

Conclusions

The results provide clear evidence that the dynamic process generating the pork cycle is, in fact, nonlinear; however, the evidence in favor of "chaos" is less conclusive. The results of this study also suggest that linear time-series models commonly used to analyze pork market dynamics (e.g., Wallis; Shonkwiler and Spreen; Bessler and Kling; Kaylen) may fail to fully capture economic dynamics. Furthermore, while GARCH error processes can account for some of the nonlinearities, the pork market appears to be characterized by other forms of nonlinear dynamics. In addition, the empirical estimates of the GP correlation dimension show that the nature of nonlinear dynamics in the pork market are complex and cannot be represented by a few state

variables. Although little a priori information is available about the exact source of these nonlinearities, fundamental asymmetries likely exist between the expansion and contraction phases of the pork cycle. Consequently, structural models are needed to provide useful information on this issue. The dynamics of entry and exit in the pork industry and its implications for supply response are promising topics for future research.

Finally, if the existence of the pork cycle requires that the cycle is not perfectly predictable (as argued in the introduction), the source of this unpredictability remains unclear. Although our results present some evidence in favor of chaos that could explain the unpredictability of the pork cycle, this evidence is not definitive. Further work is needed to assess whether chaos can help us better understand the dynamics of agricultural markets.

[Received May 1990; final revision received October 1990.]

References

- Akaike, H. "Information Theory and an Extension of the Maximum Likelihood Principle." *Second International Symposium on Information Theory*, ed. B. N. Petrov and F. Csaki. Budapest: Akademiai Kiado, 1973.
- Baumol, J. William, and Jess Benhabib. "Chaos: Significance, Mechanism, and Economic Applications." *J. Econ. Perspectives* 3(1989):77-106.
- Benettin, Giancarlo, Luigi Galgani, and Jean-Marie Strelcyn. "Kolmogorov Entropy and Numerical Experiments." *Physical Rev. A* 14(1976):2338-45.
- Bessler, David A., and John L. Kling. "Forecasting Vector Autoregression with Bayesian Priors." *Amer. J. Agr. Econ.* 68(1986):143-51.
- Bollerslev, Tim. "Generalized Autoregressive Conditional Heteroscedasticity." *J. Econometrics* 31(1986):307-27.
- Box, G. E. P., and D. A. Pierce. "Distribution of Residual Autocorrelations in Autoregressive Integrated Moving Average Time Series Models." *J. Amer. Statist. Assoc.* 64(1970):1509-26.
- Breimyer, Harold F. "Emerging Phenomenon: A Cycle in Hogs." *J. Farm Econ.* 41(1959):760-68.
- Brock, William A. "Distinguishing Random and Deterministic Systems: Abridged Version." *J. Econ. Theory* 40(1986):168-95.
- Brock, William A., and Chera L. Sayers. "Is the Business Cycle Characterized by Deterministic Chaos?" *J. Monetary Econ.* 22(1988):71-90.
- Chow, Gregory C. *Analysis and Control of Dynamic Economic Systems*. New York: John Wiley & Sons, 1975.
- Coase, R. H., and R. F. Fowler. "The Pig-Cycle in Great Britain: An Explanation." *Economica* 4(1937):55-82.
- Day, Richard H. "The Emergence of Chaos from Classical Economic Growth." *Quart. J. Econ.* 98(1983):201-13.

¹³ Recall that positive Lyapunov exponents indicate chaos.

¹⁴ The results obtained for λ_1 are consistent with those obtained by Brock; Brock and Sayers; and Frank, Genzay, and Stengos, in that the numerical estimates of λ_1 are, in general, small in magnitude.

- Ezekiel, Mordecai. "The Cobweb Theorem." *Quart. J. Econ.* 53(1938):255-80.
- Frank, Murray, R. Genzay, and T. Stengos. "International Chaos." *Eur. Econ. Rev.* 32(1988):1569-84.
- Grandmont, Jean-Michel, and P. Malgrange. "Nonlinear Economic Dynamics: Introduction." *J. Econ. Theory* 40(1986):3-12.
- Grassberger P., and I. Procaccia. "Characterization of Strange Attractors." *Physical Rev. Letters* 50(1983):448-51.
- Hao, Bai-Lin. *Chaos*. Singapore: World Scientific Publishing Co., 1984.
- Harlow, Arthur A. "The Hog Cycle and the Cobweb Theorem." *J. Farm Econ.* 42(1960):842-53.
- Hayes, Dermot J., and Andrew Schmitz. "Hog Cycles and Countercyclical Production Response." *Amer. J. Agr. Econ.* 69(1987):762-70.
- Jameson, Melvin H. "Rational Expectations and the U.S. Hog Cycle: Statistical Tests in a Linear Model." *New Directions in Econometric Modeling in Agriculture*, ed. Gordon C. Rausser. Amsterdam: North-Holland Publishing Co., 1983.
- Jelavich, Mark S. "Distributed Lag Estimation of Harmonic Motion in the Hog Market." *Amer. J. Agr. Econ.* 53(1973):223-24.
- Kaylen, Michael S. "Vector Autoregression Forecasting Models: Recent Developments Applied to the U.S. Hog Market." *Amer. J. Agr. Econ.* 70(1988):701-12.
- Larson, Arnold B. "The Hog Cycle as Harmonic Motion." *J. Farm Econ.* 46(1964):375-86.
- Ljung, G. M., and G. E. P. Box. "On a Measure of Lack of Fit in Time Series Models." *Biometrika* 65(1978):297-303.
- Lorenz, E. N. "Deterministic Non-Periodic Flow." *J. Atmospheric Sci.* 20(1963):130-41.
- Luenberger, David G. *Introduction to Dynamic Systems: Theory, Models and Applications*. New York: John Wiley & Sons, 1979.
- McLeod, A. I., and W. K. Li. "Diagnostic Checking ARMA Time-Series Models Using Squared-Residuals Autocorrelations." *J. Time Series Anal.* 4(1983):269-73.
- May, R. M. "Simple Mathematical Models with Complicated Dynamics." *Nature* 261(1976):459-67.
- Scheinkman, J., and B. LeBaron. "Nonlinear Dynamics and Stock Returns." Econ. Dep. Work. Pap. No. 181, University of Chicago, 1986.
- Shonkwiler, J. Scott, and Thomas H. Spreen. "A Dynamic Regression Model of the U.S. Hog Market." *Can. J. Agr. Econ.* 30(1982):37-48.
- Talpaz, Hovav. "Multi-Frequency Cobweb Model: Decomposition of the Hog Cycle." *Amer. J. Agr. Econ.* 56(1974):38-49.
- Takens, F. "Detecting Strange Attractors in Turbulence." *Dynamic Systems and Turbulence*, ed. D. Rand and L. Young. Lecture Notes in Mathematics No. 898. Berlin: Springer-Verlag, 1980.
- U.S. Department of Agriculture, Economic Research Service. *Livestock and Meat Statistics*. Statist. Bull. No. 552 and annual supplements. Washington DC, various issues 1973-83.
- Wallis, Kenneth F. "Multiple Time Series Analysis and the Final Form of Econometric Models." *Econometrica* 45(1977):1481-97.
- Wolf, Alan, Jack B. Swift, Harry L. Swinney, and John A. Vastano. "Determining Lyapunov Exponents from a Time Series." *Physica* 16D(1985):285-317.

State-Space Modeling of Cyclical Supply, Seasonal Demand, and Agricultural Inventories

Jeffrey H. Dorfman and Arthur Havenner

The production of certain agricultural products follows an alternate bearing pattern of large and small crops. Many agricultural products also experience seasonal shifts in demand. Such patterns present a natural opportunity for profitable inventories. This paper applies state-space modeling techniques to the problem of modeling cyclical patterns in supply and demand and presents an application to the determination of inventory levels for such agricultural products using linear quadratic control.

Key words: alternate bearing, inventories, olives, optimal control, seasonality, state-space models.

The production of certain agricultural products follows an alternate bearing pattern of large and small crops. While many agricultural products exhibit this pattern, the more famous examples are tree crops, including apples, almonds, pecans, and olives. A physiological process in the trees causes a high-yield year to be followed by a low-yield year. Such a two-year cycle will continue until some outside event, usually weather related, breaks the cycle. For example, if in a year that should have been a high yield year there is bad weather during a crucial stage of the fruit's development, a small crop may result. High yields then are expected in the year following this weather event, beginning another cyclical pattern. Because these two-year cycles are initiated by weather phenomena, all the trees in a given area tend to be on the same two-year pattern, leading to wide fluctuations in crop production for alternate bearing crops, especially those produced within a single region. Pomologists strive to reduce these yield swings and

have succeeded to some extent through fertilization, irrigation, and pruning. However, many agricultural products still have pronounced alternate bearing cycles.

The demand for many agricultural products also displays cyclical patterns related to seasonal changes in taste. Even with the recent availability of imported products virtually year-round, consumers prefer certain goods more at particular times of year. Dummy variables are too abrupt to smoothly track seasonal fluctuations in the demand schedule. A more general method of tracking such cyclical patterns in supply and demand is desirable.

This paper presents several new methods for analyzing cyclical products. It is shown that state-space models can be used successfully to model cyclical patterns in both supply and demand. Then, we demonstrate how state-space models can be inserted into the familiar linear quadratic control framework, allowing the supply and demand estimates to be employed in the determination of optimal inventory levels for such products. This will be done by applying techniques of econometrics, linear systems theory, and optimal control with an empirical example based on the California canned olive industry.

The first part of the paper briefly describes the linear systems theory approach to constructing an approximate state-space model. The second and third parts of the paper present the olive supply and demand functions which are used to forecast future production and demand for use in making inventory decisions. State-space

Jeffrey H. Dorfman is an assistant professor, Department of Agricultural Economics, University of Georgia; Arthur Havenner is a professor, Department of Agricultural Economics, University of California, Davis.

This paper is Faculty Series Paper No. 90-21 of the Division of Agricultural Economics, University of Georgia, and Giannini Foundation Paper No. 959. The research was partially supported by a Giannini Foundation grant and by Hatch Project No. H-605 of the Georgia Agricultural Experiment Station.

The authors wish to thank Desmond Jolly for helpful comments and invaluable assistance in obtaining data. The paper was also improved by comments from Masanao Aoki, Jean-Paul Chavas, Roger Craine, Ben French, Gordon King, and two anonymous referees. Any remaining errors are, as usual, the fault of the authors.

models are used in both cases. Finally, the fourth section develops the optimal control model, employing the linear quadratic framework popularized in economics by Chow. The control model produces recommendations on both absolute price levels and relative prices among the different grades.

Linear Systems Theory Estimates of State Space Models

Because the supply and demand curves of canned olives will be modeled using a particular state-space time-series procedure, the estimators of this state-space system are developed and briefly discussed below. This approach, a version of multivariate time-series analysis, emphasizes dynamic rather than static relations between variables. It uses results from the engineering literature on dynamic approximations to determine the model specification most consistent with the observed data in a formal approximation sense and has produced accurate forecasts in other contexts; see Aoki and Havenner for a survey with applications. Accurate forecasts are important in setting inventory levels with alternate bearing crops because the solution must look forward over two crop years and, using the monthly data employed here, twenty-four months of demand.

The state-space model consists of two matrix equations: the state equation,

$$(1a) \quad z_{t+1|t} = Az_{t|t-1} + Be_t,$$

and the observation equation,

$$(1b) \quad y_t = Cz_{t|t-1} + e_t,$$

where y_t is an $(m \times 1)$ vector of observations on the series to be modeled, e_t is an $(m \times 1)$ vector of stochastic innovations assumed to be i.i.d. with a zero mean, and $z_{t|t-1}$ is an $(n \times 1)$ vector of conditional means of the states, or minimal sufficient statistics for the past history of the series y_t . The subscripts on $z_{t|t-1}$ denote a conditional expectation of the state for time period t given the information set of time period $t-1$. The matrices A , B , and C are parameters to be estimated.

The innovations e_t appear in both equations. Because the states are defined to be minimally sufficient statistics for the series y_t , when a new observation occurs the states must be updated. The only new information in y_t is contained in

the innovation e_t . The matrix B serves to combine this new information into the states in a manner such that the states remain minimal sufficient statistics for the series y_t as long as the innovations are assumed to come from a distribution with the linear inheritance property (such as the normal distribution). For a detailed description of the relationship between state-space models and ARIMA models, see Aoki and Havenner.

Beginning the model estimation process, let the upper bound on the number of lags needed to model y_t be p and the desired forecast horizon be h . Define the $(mf \times 1)$ and $(mp \times 1)$ future and past data vectors as $y_t^+ = (y_t', y_{t+1}', \dots, y_{t+f-1}')$ and $(y_t^- = (y_t', y_{t-1}', \dots, y_{t-p+1}')$. The autocovariance matrix that allows the prediction of y_t^+ from y_{t-1}^- can now be defined as $H = E(y_t^+ y_{t-1}^-')$ and estimated as

$$(2) \quad \hat{H} = \begin{bmatrix} \hat{\Gamma}_1 & \hat{\Gamma}_2 & \hat{\Gamma}_3 & \dots & \hat{\Gamma}_p \\ \hat{\Gamma}_2 & \hat{\Gamma}_3 & \hat{\Gamma}_4 & \dots & \hat{\Gamma}_{p+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \hat{\Gamma}_f & \hat{\Gamma}_{f+1} & \hat{\Gamma}_{f+2} & \dots & \hat{\Gamma}_{f+p-1} \end{bmatrix},$$

where each $\hat{\Gamma}_i$ is estimated by the standard, maximum likelihood sample autocovariance formula

$$(3) \quad \hat{\Gamma}_i = T^{-1} \sum_{t=1}^{T-i} y_{t+i} y_t' \quad i = 0, 1, \dots, p+f,$$

where T is the sample size. The $\hat{\Gamma}_0$ and $\hat{\Gamma}_{f+p}$ matrices do not appear in \hat{H} , but are used later. The choice of the letter H for the autocovariance matrix is the result of its being a Hankel matrix—a block band counterdiagonal matrix.

In order to estimate the model coefficients, two matrix decompositions of the Hankel matrix will be employed. The singular value decomposition of \hat{H} is

$$(4) \quad \hat{H} = \hat{U} \hat{\Sigma} \hat{V},$$

with $\hat{U}'\hat{U}$ and $\hat{V}'\hat{V}$ identity matrices and $\hat{\Sigma}$ a rectangular matrix with the singular values along the "diagonal" beginning at the top left corner and zeros elsewhere. When estimating a model with n states, only the n -largest singular values are used. Setting the smallest singular values to zero constitutes the set of exclusion restrictions used to identify the model, equivalent to the restrictions on which lagged observations and errors have zero coefficients in an ARMA model. Denote the approximate Hankel matrix, created by replacing $\hat{\Sigma}$ with a matrix of only the n -largest

est singular values (with the other singular values set to zero), by \hat{H}_n , and the singular vector decomposition of \hat{H}_n by

$$(5) \quad \hat{H}_n = \hat{U}_n \hat{\Sigma}_n \hat{V}_n.$$

In order to derive estimators of the coefficients it is important to recognize the following pattern in the individual autocovariance submatrices of H ,

$$(6) \quad \Gamma_i = E y_{t+i} y_t' = C A^{i-1} \Omega,$$

where $\Omega = E(z_{t+1} y_t')$. This relation can be established using the model in (1) and the above definition of Ω . Recognition of this pattern allows a factorization of the Hankel matrix into its observability and controllability matrices, O and K .

$$(7) \quad H = OK = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{p-1} \end{bmatrix} [\Omega \ A \Omega \ \dots \ A^{p-1} \Omega].$$

The two different factorizations of the Hankel matrix can be equated to give

$$(8) \quad \hat{H}_n = \hat{U}_n \hat{\Sigma}_n \hat{V}_n = \hat{O}_n \hat{K}_n.$$

Equating these factorizations specifies the model and allows a solution for estimators. Let \hat{H}_1 and $\hat{H}_{\cdot 1}$ represent the first m rows and the first m columns of \hat{H} , and note the following relationships from the first block row and block column of (7) above:

$$(9) \quad H_{1\cdot} = [\Gamma_1 \ \Gamma_2 \ \Gamma_3 \ \dots \ \Gamma_p] = CK,$$

$$(10) \quad H'_{\cdot 1} = [\Gamma_1' \ \Gamma_2' \ \Gamma_3' \ \dots \ \Gamma_p'] = \Omega' O'.$$

These two equations constitute the first-order conditions for the estimators of C and Ω . The estimates of C and Ω are therefore given by

$$(11) \quad \hat{C} = \hat{H}_1 \hat{K}_n^-, \text{ and}$$

$$(12) \quad \hat{\Omega} = \hat{O}_n^- \hat{H}_{\cdot 1},$$

where \hat{K}_n^- and \hat{O}_n^- are generalized inverses defined as

$$(13) \quad \hat{K}_n^- = \hat{V}_n (\hat{\Sigma}_n)^{-1/2}, \text{ and}$$

$$(14) \quad \hat{O}_n^- = (\hat{\Sigma}_n)^{-1/2} \hat{U}_n'.$$

To estimate A , it is necessary to introduce a time-shifted autocovariance matrix, \tilde{H} , defined as

$$(15) \quad \tilde{H} = \begin{bmatrix} \Gamma_2 & \Gamma_3 & \Gamma_4 & \dots & \Gamma_{p+1} \\ \Gamma_3 & \Gamma_4 & \Gamma_5 & \dots & \Gamma_{p+2} \\ \vdots & \vdots & \vdots & \text{xxx} & \vdots \\ \Gamma_{f+1} & \Gamma_{f+2} & \Gamma_{f+3} & \dots & \Gamma_{f+p} \end{bmatrix} = OAK.$$

The left arrow notation is motivated by interpreting \tilde{H} as a Hankel matrix with the blocks of H shifted left one position and then the last column filled on the right. The second equality in (15) comes from placing A in the middle of (7) and using the relations in (6). The estimate of A is derived by applying the generalized inverses of \hat{K}_n and \hat{O}_n again, yielding

$$(16) \quad \hat{A} = \hat{O}_n^- \hat{\tilde{H}} \hat{K}_n^-.$$

The generalized inverses used for O and K determine the scale of the states and, thus, the coefficients. They were chosen so that $\hat{K}_n \hat{K}_n' = \hat{O}_n' \hat{O}_n = \hat{\Sigma}_n$. This scaling is known as the balanced representation, resulting in the coefficients in \hat{A} , $\hat{\Omega}$, and \hat{C} being uniquely identified. While other scalings could have been selected, by calculating \hat{A} , $\hat{\Omega}$, and \hat{C} from the n nonzero singular values in $\hat{\Sigma}_n$ and the orthogonal singular vectors of \hat{U}_n and \hat{V}_n subject to the balanced representation, the estimates possess the property referred to as being "strictly nested." The estimates are called "strictly nested" because the leading principal submatrices of \hat{A} , $\hat{\Omega}$, and \hat{C} contain the exact numerical values that would be produced by the model if it were estimated with the correspondingly smaller number of nonzero singular values. This result occurs because these estimates are composed of parts of scaled singular vectors (which are, of course, orthogonal). Also, since the singular values are included in decreasing order, any states left out (by too small an n) are less important than each of the included states. Thus, if additional states (associated with additional singular values) are included incorrectly in the model, all coefficients of the true model are estimated consistently and, in fact, are numerically identical to the estimates that would be obtained from the correct specification (hence the adjective "strictly"). Alternatively, if states (and associated singular values) of the true model are excluded from the estimated model, the coefficients that are included are again consistently estimated and the included states are the most important ones in the determination of the series being modeled. This has important consequences in model specification. For a detailed

proof of the strict nesting property see Havenner and Criddle.

With estimates of A , C , and Ω in hand, we can develop an estimate of the Kalman filter matrix, B . This is done by solving a system of three matrix equations involving four covariances: the state covariance matrix, $\Xi = E(z_{t+1}|z_{t+1}|')$; the error covariance matrix, $\Psi = E(e_t e_t')$; the covariance of the states with the series, $\Omega = E(z_{t+1}|y_t')$; and the unconditional data covariance matrix, $\Gamma_0 = E(y_t y_t')$. In presenting these equations in as simple a manner as possible, use has been made of the orthogonality of the current states, z_{t-1} , and the innovations, e_t . The first two equations are derived by taking the expectations of equations (1a) and (1b) postmultiplied by their own transposes:

$$(17) \quad \Xi = A\Xi A' + B\Psi B',$$

$$(18) \quad \Gamma_0 = C\Xi C' + \Psi.$$

Taking the expectation of the state equation (1a) postmultiplied by the transpose of the observation equation (1b) gives the third matrix equation,

$$(19) \quad \Omega = A\Xi C' + B\Psi.$$

Based on the previously developed estimators of A , C , Γ_0 , and Ω estimates of B , Ξ , and Ψ can be obtained by solving (17) through (19) simultaneously. Methods for solving equations (17) through (19) are found in Aoki and Havenner, and in Aoki.

With the estimates of A , B , and C that have now been derived an estimate of the initial state vector z_0 can be obtained by backcasting; for details see Aoki. Given this conditional estimate of the initial states of the system, the model given by (1) can be solved for in-sample and out-of-sample forecasts.

The Supply Model

Olive trees mature slowly and live for a long time. A tree begins bearing a crop in its seventh or eighth year, reaches full maturity at about twenty, and often produces at a profitable level up to age fifty. In the United States, commercial olive production occurs only in California. Olive bloom and fruit set occur in May, so rain does not generally present a problem with pollination and is therefore not an important factor in the determination of supply. The harvest of olives in California begins in September and ends in October or November, depending on weather

and variety. Recent acreage changes have been slightly downward (2%–3% per year), which seems to be a correction for an earlier too rapid expansion (Jolly).

The production of olives in California has varied from 5.4 to 14.3 million cases in the last ten years (California Olive Committee). Year-to-year variation has exceeded 100%. This variation is partly caused by olives being an alternate bearing crop. Pomologists associate this phenomenon with the heavy drain on the trees' nutrients during the high yield year and the subsequent small production of new fruiting wood for the next bloom period (Hartman and Opitz). This cycle within the stochastic nature of the supply of olives can be exploited for more accurate forecasting of future supplies. The main product of California olives is canned, ripe olives, with over 90% of the crop being canned. These canned olives are differentiated by size.

The supply model is concerned with predicting the supply of canned California olives in five size classes. The long lag between planting and production allows the supply to be modeled without acreage as an exogenous variable because it changed very little in the time studied and cannot change much over a two-year planning horizon. Further, supply will be modeled without price as an exogenous variable because, while the price level affects profitability and, therefore, the number of acres of olives, price has little effect on the current year's supply. As long as price is high enough to cover harvesting costs, producers will harvest their crop and cannot increase production (except, perhaps, by reducing waste) in response to a high current price.

The supply model was formulated with the five main size classes of canned olives used by the California olive industry (small, medium, large, jumbo, and colossal). Thus, supply is modeled as the joint production of five varieties of canned olives. While this ignores those olives produced but not canned and some olives that fall into a few other categories, such a formulation captures about 80% of the olives produced in California.

Structuring the optimal control model around these size classes provides two advantages. On the demand side, prices of canned olives vary by size class. By not combining all the sizes into one single product, an aggregation problem is avoided and information is preserved. On the supply side, more accuracy can be achieved without overtaxing a short data series. The supply model has a very short time series—nine annual observations. The California olive in-

dustry experienced such large growth in the early and mid-1970s that to go farther back in time would introduce structural change to the system.

With other time-series methods, this short time series could signal the end of the empirical analysis; but, with a state-space model, the shortage of data can be compensated for by the inclusion of multiple series driven by shared states. The data support our a priori conjecture that the states that determine the supply of one size of olives will be the same as, or closely related to, the states that determine the supply of other sizes, although the effects on the various sizes differ. In the empirical approximation of the model, having five series allows a higher-order determinantal polynomial to be fit to each series without utilizing the data from each series too intensively. This is because the order of the determinantal polynomial is bounded by the product of the number of series, m , and the number of lagged periods in the y_t vector, p . As the lag parameter p increases, fewer observations are available to calculate the additional estimated autocovariances [see (4)]. Therefore, the five series employ more observations in approximating the Hankel matrix (potentially resulting in greater accuracy) while not restricting the model to a specification with insufficient flexibility.

The supply model in state-space form is

$$(20) \quad z_{t+1|t} = A_s z_{t|t-1} + B_s e_t$$

$$(21) \quad y_t = C_s z_{t|t-1} + e_t$$

where y_t is a (5×1) vector of the supply of the five sizes of canned olives in year t and the s subscripts differentiate these matrices from those of the demand model which follows. The model was estimated using the years 1977 through 1985. The series were all centered and divided by standard deviations to have zero means and un-

conditional variances equal to one. This makes the mean-squared errors comparable across series. To keep the determinantal polynomial equation to a suitably low order for an annual time series of olive production, the lag parameter, p , was set to one. This allows up to a fifth-degree determinantal polynomial. It also preserves observations in the estimation of the autocovariance matrices.

The best approximation occurred with two states ($\hat{n} = 2$). The mean-squared errors for the predictions of the five series ranged from .404 to .732, which are fairly good—although not outstanding—when compared to the series' unconditional variances of unity. The estimates of the A_s , B_s , and C_s matrices are given in table 1, along with the mean-squared errors and R^2 of each of the five series. The dynamics matrix A_s has two real eigenvalues, -0.588 and 0.191 . The absolute values of both roots are less than one, implying that the system is stable. The negative eigenvalue captures the alternate bearing phenomenon by introducing a "sawtoothed" oscillation to the supply model's dynamics. Because each eigenvalue of A_s has a modulus less than one, the alternate bearing cycle would die out if it were not restarted by new exogenous weather shocks.

The Demand Model

Because California canned olives have few, if any, close complements or substitutes, the demand side of the industry also will be modeled as a state-space system with the addition of a structural equation. This allows the emphasis of the modeling to be placed on the dynamics of the system rather than on cross relations with other variables. For a product without any strong cross relations (elasticities of substitutions close

Table 1. Supply Model Parameters

	Small	Medium	Large	Jumbo	Colossal
MSE	.436	.404	.546	.732	.508
R^2	.707	.768	.570	.451	.545
A_s	-.487 -.264	-.259 .0906			
B_s	2.11 .473	-1.77 -.00804	.986 -1.97	.089 1.37	-.950 .162
C_s'	-.772 .160	-.859 .038	-.343 -.320	-.531 -.105	.0174 -.536

to zero), this is advantageous. The five sizes of California olives will again be treated as separate time series. For the demand system, though, the period of observation is one month, not a year as in the supply model. This allows seasonal cycles in demand to be traced. Such cycles are very important in olives, which have two high demand and two low demand seasons per year. Thus, a vector of demand for olive products will be estimated and forecast.

Because prices are the control variables in the optimal control model (see below), demand will be estimated with sales levels as the dependent variables and prices as an exogenous variable. The presence of exogenous (known and observable) variables, such as prices, requires that the state-space system be augmented with a regression equation in order to remove the time-varying deterministic mean ($\alpha + \beta p_t$) from the series of quantities demanded, thus rendering the resulting series stationary.

The demand system can be written as

$$(22) \quad z_{t+1|t} = Az_{t|t-1} + Be_t$$

$$(23) \quad u_t = Cz_{t|t-1} + e_t$$

$$(24) \quad q_t = \alpha + \beta p_t + u_t,$$

where q_t is a (5×1) vector of the quantity sold in month t and p_t is a (5×1) vector of the wholesale prices of the five sizes of olives. Obviously, the state vector $z_{t|t-1}$ and the innovations e_t differ from those in the supply model. This demand system was estimated using a monthly series from 1983:7 to 1987:6, saving 1987:7 to 1988:6 for out-of-sample validation. Although no prices of other goods enter the demand system, the prices of all five olive size classes are in each size's demand function. This is because there are extensive possibilities for substitutability and complementarity in demand between olive size classes.

Equations (22) through (24) were estimated by an iterative procedure. First, equation (24) was estimated by least squares. Then, using the resulting residuals, \hat{u}_t , as the observation series, (22) and (23) were estimated by the process outlined above. This process allows the construction of an updated \hat{u}_t series using the forecasts of the state-space model. The state-space model's forecasts were then used to generate a q^* series where $q_t^* = q_t - \hat{u}_t$, and \hat{u}_t is now the forecasts from (22) and (23). The structural parameters α and β were then reestimated using q^* instead of q and the process continued until the parameters converged.

The results of the estimation of the demand system in (22) through (24) are presented in table 2. The sales quantities are measured in thousands of cases. In table 2, B has been multiplied by 10^3 in order to simplify the table. Further, the mean-squared errors are normalized by each series' unconditional variance to provide a natural scale for these statistics. The MSE and R^2 statistics show that for all the size classes except the colossal a reasonable reduction in the variation of the series was obtained. The magnitude of the off-diagonal elements of β indicate that cross relations between size classes are indeed important. In fact, complementarity exists among several of the sizes; in particular, the two smallest sizes (small and medium) are complements to the three larger sizes (as can be seen by the negative terms below the diagonal in the first two columns of β). This is not surprising given the large sales of canned olive to the food service industry, where olives are used in salads and other dishes in either chopped or sliced form and only a few of the larger sized olives are employed as a decoration. While there is one positive own-price effect (for the large class), this coefficient is not significantly different from zero.

Model validation was done using a procedure from Henriksson and Merton. This procedure measures the confidence level of the hypothesis that the forecasts contain information value. This is equivalent to testing if the model can forecast the direction of revision in the series. The results of testing all five size classes together were confidence levels of 0.997 in-sample and 0.605 for the twelve out-of-sample observations. For further details see Dorfman.

The Optimal Control Model

The optimal control model is formulated as if the Olive Marketing Board could set wholesale prices for the industry. The Olive Marketing Board is comprised of grower, processor, consumer, academic, and government representatives. It is governed by a set of regulations approved by its member growers and processors under laws set up from the 1922 Capper-Volstead Act onward. While the Olive Marketing Board does not currently have the power to set prices, such power could be bestowed upon it by the vote of a suitable majority of the growers and processors. Seven processors buy essentially the entire crop of California olives from the growers, process the olives, and sell the resulting products to either food service compa-

Table 2. Demand Model Parameters

	Small	Medium	Large	Jumbo	Colossal
<i>MSE</i>	.580	.524	.615	.645	.882
<i>R</i> ²	.416	.476	.385	.355	.118
<i>A</i>	.829 -.600	.597 .290			
<i>B</i> ^a	2.25 -3.52	-.710 2.56	1.83 .860	7.21 17.91	1.86 16.64
<i>C'</i>	-10.99 13.80	-11.11 32.15	-7.96 28.43	-5.97 13.08	-.733 4.55
α'	337.4	53.34	522.4	125.4	139.3
β	-38.50 -85.00 -25.75 -5.93 -2.39	3.20 -54.76 -16.21 -12.97 -9.80	-8.08 66.08 24.05 13.86 -4.42	-12.73 77.68 61.59 -1.42 .924	15.55 12.72 38.41 -2.97 -1.13

^a *B* is presented scaled by 10³.

nies, other food processors, or retailers. Thus, the number of firms that would be involved in an inventory control program is a manageable number.

The objective function proposed for the inventory optimal control model is inspired by Vinod. He refers to models with objective functions of this type as "eclectic," implying that they attempt to mimic decision makers with multiple objectives. We propose three objectives for the regulation of olive inventories, with weights that determine their relative importance to the decision makers.

A major motivation for inventory control is to attempt to improve profit by more efficient temporal allocation of the available supply. Hence, the first term of the objective function is discounted revenues. The costs of production can be safely ignored here because the date of sale does not affect these costs. Storage costs are accounted for by adjusting the discount rate employed to reflect the percentage of the price that represents the cost of storing the canned olives.

The second objective is price stability. Both Massell and Samuelson have shown that the sum of producer and consumer surplus is higher under stable prices than instable prices. The results of Waugh, showing that consumers should favor price instability, and Oi, demonstrating a similar result for producers, result from a form of price instability that causes a more than offsetting loss to the producers or consumers, respec-

tively. When both sides of the market are considered simultaneously, price stability is welfare increasing.

In the work most familiar to agricultural economists, Newbery and Stiglitz caution against the use of consumer surplus to measure consumer welfare under various price stabilization schemes, pointing out that the Hicksian demand curve is the appropriate one to examine. Newbery and Stiglitz show that at least some degree of price stabilization, accomplished by the holding of inventories, is beneficial given reasonably low storage costs. In particular, when such stabilization does not have a significant effect on supply, as would be the case for olives with a highly price inelastic short-run supply function, Newbery and Stiglitz demonstrate that price stabilization has a net positive benefit. Craine also shows that in a general equilibrium framework stable prices are welfare increasing.

Because the marketing board represents growers and processors, this price stability term makes sense. This term also serves to discourage the regulators from "overmanaging." A term representing consumers' desire for low prices is unnecessary since the demand curves enter the optimization problem as a set of constraints.

The third objective is a desired sales level. Sales are encouraged to approximately match target levels, with a penalty for squared deviations of sales from the targets. This specification is not equivalent to a feasibility constraint (which is included in the constraint set, not in

the objective function). A feasibility constraint simply assures that sales do not exceed the sum of production and carry-in inventories.

The industry has been building up substantial inventories over the past several years with current inventory levels of approximately one-half of annual production. If storage levels continue to climb, the growers are certain to find themselves in an increasingly unfavorable negotiating position relative to the processors. The examination of these storage levels was the motivation for including desired sales levels in the objective function.

Defining ρ as the discount factor inclusive of both normal opportunity costs and storage costs, the inventory control problem is to

$$(25) \quad \min_{\{p\}} \mathcal{L} = - \sum_{t=1}^T \rho^t p'_t q_t + \sum_{t=1}^T (p_t - p_{t-1})' R_t (p_t - p_{t-1}) \\ + \left(\sum_{t=1}^T q_t - d^* \right)' S \left(\sum_{t=1}^T q_t - d^* \right),$$

subject to

$$(26) \quad z_{t+1|t} = Az_{t|t-1} + Be_t,$$

$$(27) \quad u_t = Cz_{t|t-1} + e_t,$$

$$(28) \quad q_t = \alpha + \beta p_t + u_t, \text{ and}$$

$$(29) \quad \sum_{t=1}^T q_t \leq s + k.$$

In the criterion above, the matrices R_t and S are specified by the decision makers to match their objectives, the vector d^* is the goal chosen for sales over the model period, s is a vector of production over the model period, and k is a vector of carry-in stocks. The subscript t refers to monthly time periods and T is set at 24 for the empirical solution of the model. The choice of a two-year period is based on the two-year alternate bearing cycle that olive trees follow. This natural physiological phenomenon makes one complete alternate bearing cycle an obvious choice for inventory planning horizon. Thus, the demand system of equations (26) through (28) serves as the equations of motion, while (29) is the set of feasibility constraints limiting sales to production plus carry-in inventories.

The discount parameter ρ reflects both the actual storage costs of the canned olives and the opportunity cost of waiting an additional month to sell them. The interest rate for one-month certificates of deposit of \$100,000 and over was employed. The control problem is solved for the

two-year period 1985–86. For the first week of January 1985, the interest rate was 8.1%. A reasonable estimate of storage costs was obtained from industry sources and set at 0.5% per month of the price of olives. This constant percentage markup for storage costs is reasonable because the storage facilities are sunk costs and the only other nonnegligible cost is the interest on the loans that processors use to finance themselves until the olives are sold. This can be added to the interest on the money that would have been earned by selling the olives to arrive at a total discount factor. Thus, for the 1985–86 period, the discount factor $\rho = 1/(1 + (.081/12) + .005) = 0.988$.

Because it is the relative weights of the three

terms of the objective function that determine the result, not their absolute values, the weight on the revenue term is normalized to unity. The matrices R_t and S are the relative weights for the price and feasibility terms, respectively. While they need not be diagonal in general, R_t and S will be assumed diagonal here.

The inventory control problem is solved as a deterministic closed-loop optimal control problem. It is simplest to rewrite the problem in terms of a model augmented to first order with a quadratic loss function. This can be done, paying careful attention to the dynamics embodied in the Kalman filter matrix B and to which innovations are observable in each time period. If (27) and (28) are combined and the state equation is backdated one time period, the result is a seven-equation demand model written in matrix form as

$$(30) \quad z_{t|t-1} = Az_{t-1|t-2} + Be_{t-1}$$

$$(31) \quad q_t = \alpha + \beta p_t + Cz_{t|t-1} + e_t.$$

Next, backdate (31) one period and solve for the innovation e_{t-1} in terms of q , p , and z . Substitute this expression into the right-hand side of (30) to replace the known innovation e_{t-1} . This results in equation (32) below. Then, replace the $z_{t|t-1}$ in (31) with the expression on the right-hand side of (32) and replace the unknown innovation e_t with its zero expectation. This substitution gives a new observation equation shown in (33). The result is the state-space model written in a deterministic augmented first-order form

with the dynamics of the Kalman filter preserved,

$$(32) \quad \begin{aligned} z_{t|t-1} &= Az_{t-1|t-2} + B(q_{t-1} - \alpha - \beta p_{t-1} - Cz_{t-1|t-2}) \\ &= Bq_{t-1} + (A - BC)z_{t-1|t-2} - B\alpha - B\beta p_{t-1} \end{aligned}$$

$$(33) \quad \begin{aligned} q_t &= \alpha + \beta p_t + C[Az_{t-1|t-2} + B(q_{t-1} - \alpha - \beta p_{t-1} - Cz_{t-1|t-2})] \\ &= CBq_{t-1} + C(A - BC)z_{t-1|t-2} + (I - CB)\alpha + \beta p_t - CB\beta p_{t-1}. \end{aligned}$$

The augmentation of the model necessary to write the problem with a quadratic loss function and a first-order model is straightforward. Three new variables are defined in the augmentation to first order. The variables x and v are both renamings of the price variable: $v_t \equiv p_t$ and $x_t \equiv p_{t-1}$. The variable d_t is defined by $d_t \equiv (\sum_{s=1}^t q_s - d^*)$, remembering that d^* is the desired sales over the planning horizon. The augmented model is

$$(34) \quad \bar{y}_t = \begin{bmatrix} q_t \\ z_{t|t-1} \\ v_t \\ x_t \\ d_t \end{bmatrix} = \begin{bmatrix} CB & C(A - BC) & -CB\beta & 0 & 0 \\ B & (A - BC) & -B\beta & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 \\ CB & C(A - BC) & -CB\beta & 0 & I \end{bmatrix} \bar{y}_{t-1} + \begin{bmatrix} (I - CB)\alpha \\ -B\alpha \\ 0 \\ 0 \\ (I - CB)\alpha \end{bmatrix} + \begin{bmatrix} \beta \\ 0 \\ I \\ 0 \\ \beta \end{bmatrix} p_t.$$

To achieve this definition, the negative of the vector d^* is inserted into the elements of \bar{y}_0 corresponding to d_0 . Initially, d^* is set equal to the expected production over the two-year period. This is simply the most recent year's production plus the estimate of the next year's production from the supply model in (20) and (21). At time period $t = 10$, the second year's actual production becomes known, although it has not yet been processed. Because new information is now available, d^* is then revised by making an adjustment to the elements of \bar{y}_{10} in order to incorporate the new information concerning production levels. For this example, the desired sales level is simply total production. In general, d^* need not equal production. In cases where undesirably large stocks are left over from earlier years, d^* might exceed production in order to reduce stocks.

Noting that all the matrices in the model are time invariant, define new notation for (34) as

$$(35) \quad \bar{y}_t = \bar{A}\bar{y}_{t-1} + \bar{b} + \bar{C}p_t.$$

The loss function can be written in a single quadratic form with a new penalty matrix, Π_t , given by

$$(36) \quad \Pi_t = \begin{bmatrix} 0 & 0 & -.5\rho I & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ -.5\rho I & 0 & R_t & -R_t & 0 \\ 0 & 0 & -R_t & R_t & 0 \\ 0 & 0 & 0 & 0 & S_t \end{bmatrix},$$

where S_t is a null matrix for $t = 1, 2, \dots, T - 1$, and $S_T = S$ where S is the penalty matrix from

(25). The matrices R_t are identical in all periods except two, the first and the eleventh. R_1 is set equal to the null matrix to allow the optimal control model to select the first period's prices p_1 without any penalty for deviations from the initial price levels p_0 . R_{11} is set to one-tenth its value in other periods to allow for the prices to adjust to the added information concerning the actual production levels. It was decided not to set R_{11} equal to the null matrix, thus forcing a

smoother adjustment to the new information.

The optimal control problem can now be written as

$$(37) \quad \min_{\{p\}} \mathcal{L} = \frac{1}{2} \sum_{t=1}^T \bar{y}_t' \Pi_t \bar{y}_t,$$

subject to

$$(38) \quad \bar{y}_t = \bar{A}\bar{y}_{t-1} + \bar{b} + \bar{C}p_t,$$

$$(39) \quad \sum_{t=1}^T q_t \leq s + k.$$

The problem has been transformed into the well-known linear-quadratic deterministic optimal control problem. Given estimates of \bar{A} , \bar{b} , and \bar{C} from the demand system, and a set of penalty matrices, Π_t , the problem given by (37) through (39) can be solved by forming a Lagrangian, taking the derivatives with respect to \bar{y}_t , p_t , and the Lagrange multipliers, and setting them equal to zero. The resulting first-order conditions can be solved to yield a set of Riccati equations which allow for a recursive solution for the parameters of the optimal feedback rule. See Chow for further details. Skipping the intermediary algebra, the feedback rule for choosing the optimal policy instruments, \bar{p}_t , is given by the following equations:

$$(40) \quad \bar{p}_t = G_t \bar{y}_{t-1} + g_t,$$

$$(41) \quad G_t = -(\bar{C}'H_t\bar{C})^{-1}\bar{C}'H_t\bar{A},$$

$$(42) \quad g_t = -(\bar{C}'H_t\bar{C})^{-1}\bar{C}'(H_t\bar{b} - h_t),$$

$$(43) \quad H_{t-1} = \Pi_{t-1} + \bar{A}'H_t(\bar{A} + \bar{C}G_t),$$

$$(44) \quad h_{t-1} = (\bar{A} + \bar{C}G_t)'(h_t - H_t\bar{b}),$$

$$(45) \quad H_T = \Pi_T, \text{ and}$$

$$(46) \quad h_T = 0,$$

where the bars on \bar{A} and \bar{C} and the subscript on H_t distinguish them from the A , C , and H matrices used in the state-space model. Equations (41) through (46) are solved recursively for a set of G_t and g_t , which then allow the optimal price vector for each period to be calculated after the past period's sales have been observed. The feasibility constraints in (39) can be ignored in the solution as they are nonbinding in all time periods.

For the inventory control model, the submatrices in Π_t are set to adjust for the relative scales of the three terms in the objective function: revenue terms are on the order of 10^3 . The selections are $R_t = \text{diag. } (50000, 50000, 50000, 50000, 50000)$, with the adjustments discussed above for the time periods $t = 1$ and $t = 11$, and $S_T = \text{diag. } (.25, .25, .25, .5, .5)$. These values were chosen so that variations in the variables consistent with historical values would not overwhelm, or be overwhelmed by, the revenue terms. In particular, the last two terms in S_T were set to twice the level of the first three because the two corresponding size classes (jumbo and colossal) have production levels significantly lower than the three other size classes. Therefore, the same absolute difference in cases represents a larger percentage difference and deserves to be treated as a larger loss (i.e., smaller misses from the target are larger in a relative sense). For the submatrices in Π_t relating to revenue, the discount parameter is set to $\rho = .988$ as specified above.

Results of the Model Solution

The results of the optimal control model solution showed that, while the model is just an approximation of actual industry behavior, the technique can produce useful policy implications. It would also be easy for an industry to

employ such a method and adapt it to their objectives.

The model took the alternate bearing phenomenon into account by choosing cross-year inventory levels that ranged from one quarter of one month to two months worth of sales for the five size classes. The cross-year inventory levels chosen differed significantly from those that occurred, but the comparison is confused by the large carry-in stocks. These stocks are not factored into the sales targets (although they could be), entering only through the feasibility constraint.

When the information on the actual production levels for the second year entered the model, it adjusted quickly. Over the two-year period, the actual production levels were higher than the forecast. Upon receiving this "news," the prices of all five size classes dropped to allow for increased sales. Prices remained at lower levels until the end of the control period, allowing the model to achieve total sales within 2% to 4% of the desired levels for all the size classes. The values for the differences between the sales and production for the two-year period for the actual industry and for the model are given in table 3. A positive number indicates that carry-in stocks were reduced, and all positive numbers are within the feasible region (the feasibility constraints are all nonbinding). Again, the large positive numbers for the actual industry in all the size classes except for the colossal indicate that the industry is reducing its "permanent" inventory levels. The failure to reduce the inventory levels of the colossal size is important because considerable stocks exist in this size class.

Another positive sign of the model's performance was that the sales levels followed a cyclical pattern, peaking in the months when the seasonal demand for olives is highest. This implies that the state-space part of the demand model is allowing the seasonal variations in demand to be tracked with reasonable accuracy. The cycle embodied by the state-space model allows for a smoother transition to these seasonal changes than the inclusion of dummy variables in the demand system.

The price levels set by the model diverged from the actual levels that occurred. The prices set by the model for the large size class were higher than the actual ones; for the medium, jumbo, and colossal sizes the prices set were lower, and for the small size class the prices set by the model were approximately the same as actually occurred. This feature of the solution is the result of the sales constraint imposed through the pen-

Table 3. Sales Minus Production for 1985–86

	Small	Medium	Large	Jumbo	Colossal
Proposed	-5.9	6.2	16.4	-2.9	-59.6
Actual	205.5	1075.9	141.5	556.2	-398.2

Note: All quantities in 1,000 cases.

ality matrix S . Examination of industry statistics shows that, for the past five years, inventory levels have been increasing steadily and now equal one-half of an average year's production. The control model did not allow the industry to increase stocks greatly, as it has often been doing. The results of the control model suggest that the industry must adjust its pricing structure to avoid excessive inventory levels. The model selected prices which vary more by size class than the actual prices, keeping prices high for the size classes with strong demand while reducing other prices to avoid increasing inventories further. For example, for the highest volume size class, large, the control model suggests about a 12% price increase; while, for the two lowest volume classes, jumbo and colossal, the control model suggests price cuts of approximately 38% and 57%, respectively. The ratios of the means of the proposed prices to the means of the actual prices are presented in table 4. Recalling the numbers in table 3 for sales minus production, the huge difference between the proposed and the actual prices for the colossal size class suggest that the industry is building up large "permanent" inventories of the colossal size class. To stop this practice, the model suggest a price cut of over 50%.

Finally, the results of the optimal control model are not sensitive to the values of the penalty matrices R and S . Experimentation with a range of values within the same order of magnitude showed no significant sensitivity of the price or sales levels to the exact values of R and S . This suggests that the answers produced by the control model are a result of the supply and demand models and the general form of the criterion function rather than being forced onto some "desired" path by the penalty matrices.

Table 4. Ratio of Mean Proposed to Mean Actual Prices 1985–86

Small	Medium	Large	Jumbo	Colossal
0.88	0.74	1.12	0.62	0.43

Conclusions

This paper has shown that state-space models can be used to predict the cyclical patterns present in the supply and demand functions of many agricultural products. State-space models are well-suited to such problems because they can model both the alternate bearing phenomenon on the supply side and any seasonal fluctuations in demand. Such techniques may prove useful to a number of agricultural industries, as many tree crops exhibit an alternate bearing pattern in their production of fruit or nuts. An augmentation of the state-space model which allows it to be analyzed in the well-known linear-quadratic control framework while properly preserving the dynamics of the system was developed. This allows the state-space models to be used in control problems with known analytical solutions, avoiding the need for complex iterative algorithms. As an example, it was shown that state-space models can be applied to the problem of selecting cross-year inventory levels for agricultural crops with alternative bearing patterns.

In the applications presented here, policy conclusions were drawn related to the industry's pricing policies. It appears that fairly large changes in both relative and absolute prices are needed if the industry wants to avoid building up a sizable permanent inventory. Such a permanent carry-over of stocks is becoming a feature of the industry. Finally, although the model was solved as a closed-loop control problem, over the twenty-four months of the control simulation there were no errors to feed back into the model. If such a model were actually employed, the results should improve since any exogenous shocks could be incorporated into the solutions.

[Received July 1989; final revision received October 1990.]

References

- Aoki, M. *State Space Modeling of Time Series*. Berlin: Springer-Verlag, 1987.

- Aoki, M., and A. Havenner. "State Space Modeling of Multiple Time Series." *Econometric Rev.* 10(1991):1-59.
- California Olive Industry Statistics. Fresno: California Olive Committee, various years.
- Chow, G. C. *Analysis and Control of Dynamic Economic Systems*. New York: John Wiley & Sons, 1975.
- Craine, R. "Risky Business: The Allocation of Capital." *J. Monetary Econ.* 23(1989):201-18.
- Dorfman, J. H. *Three Essays Involving Time Series Analysis*. Ph.D. thesis, University of California, Davis, 1989.
- Hartmann, H., and K. Opitz. *Pruning Olive Trees in California*. University of California, Berkeley, Agr. Exp. Sta. Circular No. 537, 1966.
- Havenner, A., and K. Criddle. "System Theoretic Time Series: An Application to Inventories and Prices of California Range Cattle." *Computers Math. Appl.* 17(1989):1177-87.
- Henriksson, R. D., and R. C. Merton. "On Market Timing and Investment Performance II: Statistical Procedures for Evaluating Forecasting Skills." *J. Bus.* 54(1981):513-33.
- Jolly, D. *California Olive Industry Performance and Prospects*. University of California, Davis, Coop. Extens. Serv. 1985.
- Massell, B. F. "Price Stabilization and Welfare." *Quart. J. Econ.* 83(1969):285-97.
- Newbery, D. M. G., and J. E. Stiglitz. *The Theory of Commodity Price Stabilization*. Oxford: Clarendon Press, 1981.
- Oi, W. Y. "The Desirability of Price Instability Under Perfect Competition." *Econometrica* 27(1961):58-64.
- Samuelson, P. A. "The Consumer Does Benefit from Feasible Price Stability." *Quart. J. Econ.* 86(1972):476-93.
- Vinod, H. D. "Regulatory Economics: Dynamic Optimization and Control Under Multiple Objectives." Fordham University, 1989.
- Waugh, F. W. "Does the Consumer Benefit from Price Instability." *Quart. J. Econ.* 58(1944):602-14.

Perennial Crop Supply Response: A Kalman Filter Approach

Keith C. Knapp and Kazim Konyar

A state-space model for perennial crop supply response is developed. New plantings and removals depend on the existing age structure of the crop and expected values for future prices and other exogenous variables. Acreage in individual age categories evolves depending upon existing acreage, new plantings, and removals. The Kalman filter and an iterative parameter search provide maximum-likelihood estimates of the unknown parameters and age group acreages from observed data on total acreage and production. An empirical application for alfalfa shows that existing acreage has differential impacts on new plantings and removals depending upon age.

Key words: Kalman filter, perennial crops, state space models, supply.

A significant literature has developed on the estimation of perennial crop supply response. A number of studies are based on single-equation regression models for either aggregate output, aggregate acreage, or changes in these variables. Akiyama and Trivedi provide a review. This approach is limited because decisions on new plantings and removals depend on existing acreage in the various age categories, as well as existing total acreage and expectations about future prices, yields, and other variables. In addition, the structural models used to derive the reduced-form estimating equations typically assume desired levels of acreage or output as a function of prices and a partial adjustment or stock adjustment model to determine investment or changes in acreage or output. With stable prices and other exogenous variables and an adjustment coefficient between zero and one, this implies convergence to a steady-state or long-run equilibrium, which need not be the case for perennial crops (Knapp).

Several studies have provided separate estimates of new planting and removal equations which account, at least in part, for age distribution effects. These include French and Bressler for lemons; Arak for coffee; Alston, Free-

bairn, and Quilkey for oranges; French, King, and Minami for cling peaches; Akiyama and Trivedi for tea; and Hartley, Nerlove, and Peters for rubber. However, these studies utilize data on new plantings, removals, and age distribution which are not generally available for most crops in most regions. Bellman and Hartley propose an approach utilizing "neoclassical econometrics," but no empirical estimates are given. Knapp investigates a normative model for alfalfa.

For many perennial crops there is limited or no data on new plantings area, removals, and area in individual age groups. Thus, the problem remains of econometrically estimating perennial crop supply response with aggregate acreage and production data while still accounting for age-group dynamics in a theoretically correct manner. This paper proposes the use of state-space models and the Kalman filter to solve this problem. State-space models have system dynamics with (possibly) unobservable state variables and measurement equations relating the state variables to observable variables. The Kalman filter generates optimal estimates of the state variables and their variance-covariance matrix. The perennial crop problem is reformulated here as a state-space model, and the Kalman filter is used in conjunction with an iterative search to generate parameter estimates as well as estimates of new plantings, removals, and existing acreage by age group. An empirical application to California alfalfa production is given.

Keith C. Knapp is an associate professor of resource economics, University of California, Riverside. Kazim Konyar is an agricultural economist, U.S. Department of Agriculture.

Phyllis Nash provided assistance with the graphics. The authors are also grateful to two anonymous referees for comments motivating an expansion of the original analysis.

Theory

An individual producer is considered. The producer makes new plantings and removes existing acreage to maximize profits over an infinite horizon. The model, derived from Knapp, is also a simplified version of the model in Bellman and Hartley. The following variables are defined: A_{1t} is new plantings, beginning of year t ; A_{it} is existing area of the perennial crop at the beginning of year t which will be in the i th year of production during year t if produced, $i = 2, \dots, n$; and R_{it} is the area of the perennial crop in the i th year of growth which is removed from production at the beginning of year t , $i = 2, \dots, n$. Here n is the maximum number of age categories considered. A_{1t} and R_{it} are control variables, while A_{it} , $i = 2, \dots, n$ are the state variables.

The age group dynamics are given by

$$(1) \quad A_{i,t+1} = A_{i-1,t} - R_{i-1,t},$$

where $i = 2, \dots, n$ and $R_{1t} = 0$. Profits from the perennial crop depend on the cost of new plantings and removals, annual production costs, the opportunity cost of land planted to the perennial crop, and revenue received. Present value of profits is given by

$$(2) \quad \sum_{t=1}^{\infty} \alpha^t \pi(A_{1t}, \dots, A_{nt}, R_{1t}, \dots, R_{nt}; P_t),$$

where α is the discount factor, P_t is a vector of prices and other exogenous variables, and π denotes annual profits as a function of acreage, removals, and prices. The producer's problem is to maximize (2) subject to (1) and nonnegativity conditions on the variables.

This is a dynamic programming problem. Under appropriate assumptions about π , an optimal value function J_t exists for each year which gives the present value of profits in future years under optimal operation:

$$(3) \quad J_t(A_{2t}, \dots, A_{nt}; P_t, \dots, P_{\infty}) \\ = \sum_{s=t}^{\infty} \alpha^s \pi(A_{1s}^*, \dots, A_{ns}^*, R_{1s}^*, \dots, R_{ns}^*; P_s),$$

where A_{it}^* and R_{it}^* denote optimal levels of the state and control variables; J_t is a function of the state variables and depends parametrically on prices and other exogenous variables.

Optimal values for the decision variables in each year are chosen to maximize

$$(4) \quad \pi(A_{1t}, \dots, A_{nt}, R_{1t}, \dots, R_{nt}; P_t) \\ + \alpha J_{t+1}(A_{2,t+1}, \dots, A_{n,t+1}; P_{t+1}, \dots, P_{\infty})$$

subject to (1) and the nonnegativity conditions. Optimal decision rules can then be written as

$$(5) \quad A_{1t}^* = f_1(A_{2t}, \dots, A_{nt}; P_t, \dots, P_{\infty})$$

$$(6) \quad R_{it}^* = g_i(A_{2t}, \dots, A_{nt}; P_t, \dots, P_{\infty})$$

for $i = 2, \dots, n$. Thus, the optimal values for A_{1t} and R_{it} depend on existing acreage of the crop in each age group and future prices. A theoretical model for perennial crop supply response is given by (1), (5), and (6) after specification of the appropriate functions.¹

Empirical Model

A linear model is specified for empirical estimation. Observations are assumed to be available for total acreage and production of the perennial crop and various exogenous variables believed to influence crop supply. However, new plantings, removals, and acreage by age category are not directly observable. This is the situation for many perennial crops.

New plantings are given by

$$(7) \quad A_{1t} = \sum_{j=2}^n a_j A_{jt} + \sum_{j=1}^m b_j Z_{jt} + e_{1t}^n,$$

where Z_{jt} are exogenous variables, m is the number of exogenous variables, and e_{1t}^n is a white-noise error term; $Z_t = (Z_{1t}, \dots, Z_{mt})$ includes expected values for future prices and other exogenous variables. Removals are specified as

$$(8) \quad R_{it} = \sum_{j=2}^n c_{ij} A_{jt} + \sum_{j=1}^m d_{ij} Z_{jt} + e_{it}^r,$$

where $i = 2, \dots, n$ and e_{it}^r is a white-noise error term. Total acreage during year t is then

¹ The model in this section of the paper implicitly assumes that future values of prices (and other exogenous variables) are either known with certainty or that the conditions implying certainty equivalence are met. Neither of these assumptions is likely to be completely true in practice. The implication is that the optimal decision rules for new plantings and removals [eqs. (5), (6) in the text] depend parametrically on all moments of the probability distribution for future prices and other exogenous variables and not just the expected values as implied in the text. In the ensuing empirical analysis, we follow most studies in empirical agricultural supply response analysis and consider only the expected values of price and other exogenous variables, while acknowledging that a more complete empirical analysis would include higher moments as well.

$$(9) \quad A_t = \sum_{i=1}^n (A_{it} - R_{it}) + e_t^a,$$

where e_t^a is again a white-noise error term. Equation (9) assumes that total acreage (A_t) is subject to measurement error. Total production is given by

$$(10) \quad Q_t = \sum_{i=1}^n y_{it}(A_{it} - R_{it}) + e_t^q,$$

where y_{it} is crop yield for age-category i in year t . The y_{it} are assumed to be known coefficients; e_t^q is also a white-noise error term which is independent of e_t^a , e_t^n , and e_{it}^r .

The complete empirical model is equations (1) and (7)–(10). The coefficients to be estimated are a_j , b_j , c_{ij} , d_{ij} , and the covariance matrices for the error term vectors (e_t^a , e_t^q) and (e_t^n , e_{it}^r). Given estimates of these coefficients and the expected values and variance-covariance matrix for A_{io} and R_{io} , then estimates of A_{it} and R_{it} , $t = 1, \dots, T$, can also be constructed. These estimations are carried out by recasting the model in state-space form and using a Kalman filter combined with an iterative search over the parameters. This procedure is described in the next section.

Estimation Procedure

The perennial crop empirical model defined by equations (1) and (7)–(10) is first recast in state-space form. Substituting (1) into (7) and (8) gives

$$(11) \quad A_{it} = \sum_{j=2}^n a_j(A_{j-1,t-1} - R_{j-1,t-1}) + \sum_{j=1}^m b_j Z_{jt} + e_t^n,$$

$$(12) \quad R_{it} = \sum_{j=2}^n c_{ij}(A_{j-1,t-1} - R_{j-1,t-1}) + \sum_{j=1}^m d_{ij} Z_{jt} + e_{it}^r,$$

where $i = 2, \dots, n$. In (11) and (12) the Z_{jt} are exogenous variables with known values; however, the coefficients b_j and d_{ij} are unknown parameters. In the state-space model, b_j and d_{ij} can therefore be treated as state variables with system dynamics given by

$$(13) \quad b_{jt} = b_{j,t-1} \text{ and}$$

$$(14) \quad d_{ijt} = d_{ij,t-1},$$

where $i = 2, \dots, n$ and $j = 1, \dots, m$. Mean

values and variance-covariance matrices for the initial values of b_j and d_{ij} must also be specified.

The state variables for the empirical model in state-space form are A_{it} , $i = 1, \dots, n$; R_{it} , $i = 2, \dots, n$; b_{jt} , $j = 1, \dots, m$; and d_{ijt} , $i = 2, \dots, n$, $j = 1, \dots, m$. The observed variables are A_t and Q_t , $t = 1, \dots, T$. The measurement equations are (9) and (10), while system dynamics are given by equations (11), (1), and (12)–(14). The unknown parameters to be estimated in the state-space model are a_j , c_{ij} , the variance-covariance matrices for the error terms, and the expectations and variance-covariance matrix for the state variables at $t = 0$. Let this set of unknown parameters and initial conditions be denoted U .

With these equations and variable definitions, the model is in state-space form for given values of the parameters and initial conditions in U (Judge et al.). The Kalman filter can then be used to give optimal estimates of the state-space state variables and their variance/covariance matrix for given values in U , as well as the value of the log-likelihood function. Iterations over the values of the parameters and initial conditions in U are then conducted to find maximum-likelihood estimates for all unknown coefficients. Details for the particular empirical application considered here are available from the authors upon request.

Data

The model is applied to alfalfa production in California. Alfalfa is typically planted for three to four years and then removed from production. Therefore, the empirical model is estimated with four age groups. The estimations are carried out using data from 1945–85. Except where noted, data sources are described in Konyar and Knapp.

Age-specific yields are required to specify the quantity measurement equation (10). Based on Knapp, relative yields are defined by $y_{it} = a_i y_{2t}$, $i = 1, \dots, n$ and $(a_1, \dots, a_n) = (.66, 1., .92, .84)$. There has been a significant upward trend in California alfalfa yields. A time-series regression for the period under consideration gives

$$\bar{y}_t = 4.31 + .054 * t \quad R^2 = .92, \\ (72.76) \quad (21.12)$$

where \bar{y}_t denotes the trend in average yield (tons/acre) and t -statistics are given in parentheses. The average expected yield in each period is as-

summed to equal the trend of the average yield in that period:

$$\sum_{i=1}^n y_{it}/n = \bar{y}_t,$$

which implicitly defines y_{2t} and hence y_{it} for each period t using the definitions and coefficients described above.

From the theoretical model, new plantings and removals depend on existing acreage and expected profitability of growing alfalfa. Expected profitability depends in turn on expected revenue from alfalfa minus production costs and the opportunity cost of using land for alfalfa production. Prices received for competing crops are used here as a surrogate for land opportunity costs. The new plantings and removal functions are therefore specified as

$$(15) \quad A_{it} = \sum_{j=2}^4 a_j (A_{j-1,t-1} - R_{j-1,t-1}) + b_1 + b_2 TRALF_t^* + b_3 PCINDX_t^* + b_4 CCINDX_t^* + e_t^n \text{ and}$$

$$(16) \quad R_{it} = \sum_{j=2}^4 c_{ij} (A_{j-1,t-1} - R_{j-1,t-1}) + d_{i1} + d_{i2} TRALF_t^* + PCINDX_t^* + d_{i4} CCINDX_t^* + e_t^r,$$

where $TRALF_t^*$ is the expected total revenue from alfalfa production (\$/acre), $PCINDX_t^*$ is the expected value for a production cost index, and $CCINDX_t^*$ is the expected value of an index for the prices of competing crops; $TRALF_t^*$ is further defined as the expected price of alfalfa times expected yield, where expected yield is calculated from the above regression of alfalfa yields on time.

The main empirical results presented in the paper assume naive expectations for prices and production costs. Thus, expected alfalfa prices, $PCINDX_t^*$ and $CCINDX_t^*$ just equal their actual values lagged one year. $PCINDX$ (1977 = 100) is based on U.S. Department of Agriculture (USDA) statistics, while $CCINDX$ (1975 = 1.0) is constructed by calculating a weighted average of prices received for eight California field crops and then using 1975 as the base year. A separate section also considers quasi-rational expectations in the sense of Nerlove, Grether, and Carvalho. In this case, expectations are calculated from ARIMA models for the individual time series.

Finally, the error terms in the measurement and systems equations are assumed to be mutually independent. This implies that only the variances of the individual error terms need to be estimated; the covariances are assumed to be zero.

Empirical Results

Table 1 gives the estimated coefficients for the new plantings equation (15). The negative age-group coefficients in the new plantings equation imply that, everything else equal, an increase in existing acreage results in a decrease in new plantings. This is consistent with a priori expectations based on theoretical considerations.² Also, both the magnitude (absolute value) and statistical significance of the age-group coefficients a_j decrease as j increases. Thus, existing acreage in older age categories has less influence on new plantings than does existing acreage in younger age categories.

The coefficient for $TRALF_t^*$ is positive, while the coefficients for $PCINDX_t^*$ and $CCINDX_t^*$ are negative. Thus, new plantings increase as expected gross revenues from alfalfa production

increase, as expected production costs decrease, or as the expected returns from other crops decrease. These results are also consistent with a priori considerations. The R^2 value for the new plantings equation is somewhat low (.293) suggesting that other functional forms or explanatory variables could be tried in future work. The Durbin-Watson statistic does not suggest the presence of serially correlated error terms.

Coefficient values for the removal equations (16) are also given in table 1. The age-group coefficients c_{ij} are all positive for $i = j$ and significantly larger in absolute value than the coefficients c_{ij} , $i \neq j$, by an order of magnitude or more. These results are consistent with biological considerations in growing perennial crops. In general, some fraction of existing acreage is removed each year because of poor stands, disease, or pest damage. Also, the length of time alfalfa stands are kept in the ground varies regionally due to climatic differences. For example, alfalfa stands may typically last three years in the Imperial Valley in California but four years in the Central Valley. Thus, holding prices constant, removals in age group j will be positively

² Imagine the system is initially at a steady state. If $a_j > 0$, then an increase in existing acreage over the steady-state level would increase new plantings and total acreage would continually increase over time. If $a_j < 0$, then an increase in existing acreage decreases new plantings and the system can return to a steady state.

Table 1. Estimated New Plantings and Removals Equations for the Alfalfa Supply Response Model (naive expectations)

Independent Variable	Dependent Variable			
	A_1	R_2	R_3	R_4
A_2	-.584 (.170)*** ^a	$.489 \times 10^{-1}$ ($.618 \times 10^{-4}$)***	$-.184 \times 10^{-3}$ ($.723 \times 10^{-4}$)***	$-.387 \times 10^{-2}$ ($.683 \times 10^{-4}$)***
A_3	-.308 (.156)**	$.375 \times 10^{-2}$ ($.567 \times 10^{-4}$)***	.304 ($.664 \times 10^{-4}$)***	$.146 \times 10^{-2}$ ($.628 \times 10^{-4}$)***
A_4	-.171 (.158)	$-.438 \times 10^{-2}$ ($.574 \times 10^{-4}$)***	$-.195 \times 10^{-2}$ ($.672 \times 10^{-4}$)***	.297 ($.635 \times 10^{-4}$)***
Constant	.843 (.150)***	$-.451 \times 10^{-2}$ ($.545 \times 10^{-4}$)***	$-.482 \times 10^{-2}$ ($.638 \times 10^{-4}$)***	$-.446 \times 10^{-3}$ ($.603 \times 10^{-4}$)***
TRALF*	$.729 \times 10^{-3}$ ($.487 \times 10^{-3}$)*	$-.946 \times 10^{-5}$ ($.177 \times 10^{-6}$)***	$-.118 \times 10^{-4}$ ($.207 \times 10^{-6}$)***	$-.532 \times 10^{-5}$ ($.196 \times 10^{-6}$)***
PCINDX*	$-.125 \times 10^{-2}$ ($.128 \times 10^{-2}$)	$.123 \times 10^{-4}$ ($.466 \times 10^{-6}$)***	$.187 \times 10^{-4}$ ($.546 \times 10^{-6}$)***	$.596 \times 10^{-5}$ ($.516 \times 10^{-6}$)***
CCINDX*	-.323 (.159)**	$.297 \times 10^{-2}$ ($.577 \times 10^{-4}$)***	$.372 \times 10^{-2}$ ($.676 \times 10^{-4}$)***	$.155 \times 10^{-2}$ ($.638 \times 10^{-4}$)***
R^2	.293	.999	.999	.999
D-W ^b	1.78	2.15	2.65	1.72

Note: Results are from the OLS regressions in the last iteration of the iterative KF/OLS algorithm.

^a Standard deviations of the estimated coefficients are in parentheses; single, double, and triple asterisks denote significance at the 10%, 5%, and 1% levels, respectively.

^b D-W is the Durbin-Watson statistic.

related to, and most strongly affected by, existing acreage in the same age group.

The results in table 1 also show that removals decrease as expected alfalfa gross revenues increase and increase as expected production costs and returns from competing crops increase. All coefficient values in the removal equations are significant at the 1% level, the R^2 values are quite high, and the Durbin-Watson statistics are not suggestive of serial correlation in the error terms.

Figure 1 shows estimated new plantings and removals of age categories 2 and 3 from the KF smoother recursions. Estimated new plantings exhibit considerable variability. Removals of second-year acreage are almost constant; removals of third-year acreage exhibit somewhat greater variability. Estimated values for year 4 removals (not reported) are somewhat lower on average than year 3 removals and exhibit a slightly lower range of variation. Time-series plots for acreage in age categories 2–4 (also not reported) are qualitatively similar to the time-series plot for new plantings in figure 1 after a suitable number of lags. Because removals are positive, acreage in each age category is somewhat less than acreage in the previous age category, and this also tends to dampen the observed variability over time.

The observed variables in this analysis (other than exogenous variables) are assumed to be alfalfa production and total acreage in California.

Fitted values for A_{it} and R_{it} are calculated from equations (1) and (15)–(16) with the error terms set to zero and where estimated values from the Kalman filter smoother recursions are used for the unobserved acreage and removal variables on the right-hand side of these equations. The fitted values for A_{it} and R_{it} are used to calculate fitted production and total acreage, and these are plotted in figure 2 along with actual values. This procedure is analogous to computing fitted values for ordinary least squares (OLS) regressions.

Fitted production tracks actual alfalfa production quite well, with significant deviations in only a few years. The model tracks total alfalfa acreage somewhat less well. Although the fitted results follow the overall observed trends in actual areas, there are significant deviations in several years. These results correspond with the estimated standard deviations of the error terms for the measurement equations. For production, the estimated standard deviation of e_t^q is 9.8×10^{-6} tons and the standard deviation of e_t^a is 4×10^4 acres. (Note that both standard deviations are relatively low in comparison to an average total alfalfa area of 1.1×10^6 acres and yield of 5.4 tons/acre over the period 1945–85.)

Long-run equilibrium or steady state occurs for values of A_{it} and R_{it} which, once achieved, are maintained indefinitely for given values of the exogenous variables. Long-run equilibrium

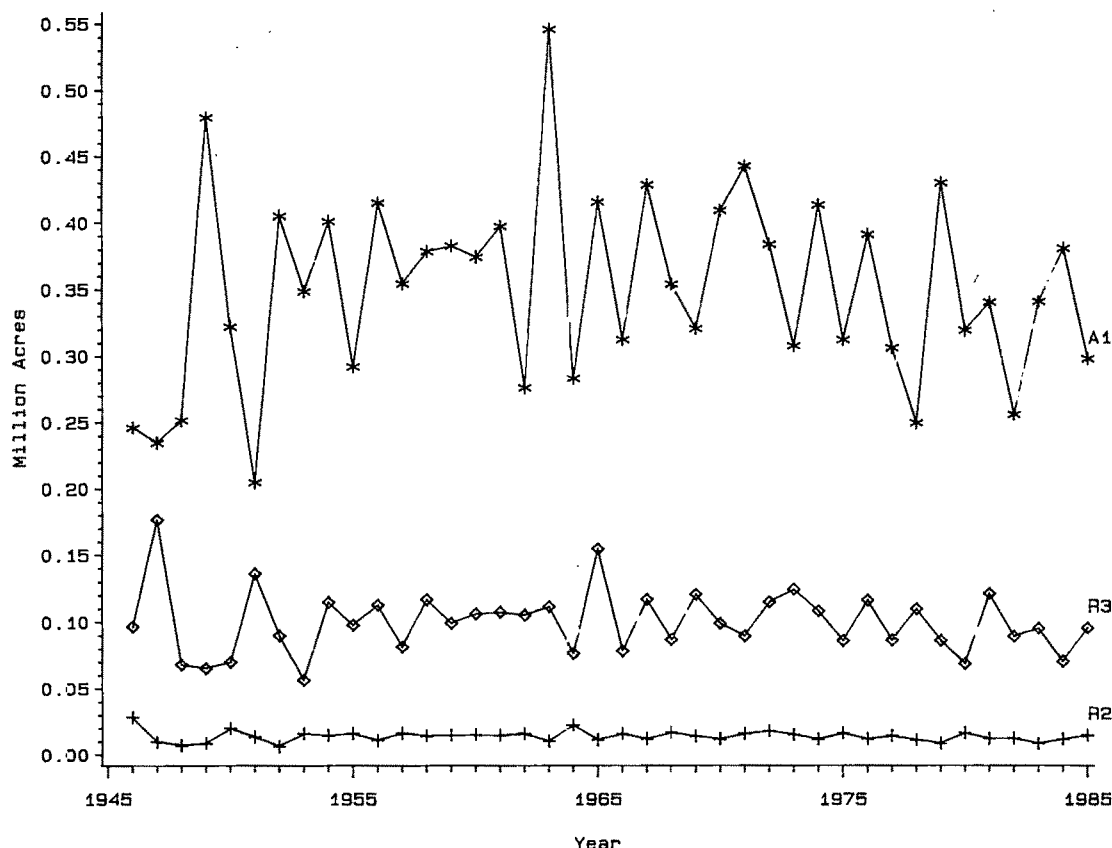


Figure 1. Time-series plots for estimated new alfalfa plantings in California (A1), and removals in age-categories 2 (R2) and 3 (R3)

values for A_{it} and R_{it} can be calculated from equations (1) and (15)–(16) after dropping the time indices. This constitutes a system of seven equations in seven variables (noting that $R_1 = 0$), which was solved using *Mathematica* (Wolfram) to obtain steady-state values of A_i and R_i as functions of the exogenous variables. Total alfalfa acreage in long-run equilibrium is then given by

$$(17) \quad A^* = 1.32 + 1.17 \times 10^{-3} \text{TRALF}^* - 1.99 \times 10^{-3} \text{PCINDX}^* - .512 \text{CCINDX}^*$$

for the estimated parameter values in table 1. Using 1983–85 average values for the exogenous variables, long-run equilibrium total acreage is estimated as 1.08 million acres which is quite close to the average of 1.1 million acres over the period 1945–85.

Price elasticities are reported in table 2 for quantity supplied in the California alfalfa market. Short-run (1 year) elasticities are calculated from equation (9) after substituting in the esti-

mated equations (15)–(16), dropping the error term, and using 1983–85 average estimated values for A_{it} from the KF smoother recursions. Long-run elasticities are calculated using equation (17). All calculations are based on 1983–85 average values for prices, yields, and acreage. The results suggest that alfalfa supply is inelastic in both the short and long run.

Forecasted alfalfa area was calculated under alternate initial conditions assuming 1983–85 average values for yield and the exogenous variables. One simulation uses estimated values from the KF smoother recursions to specify initial (1985) values for A_{it} and R_{it} . Since 1985, total alfalfa acreage is quite close to the long-run equilibrium acreage calculated earlier, the calculated time path is relatively constant; there are some minor fluctuations in the early years because individual age groups are not at long-run equilibrium levels. To further explore the dynamic behavior of the estimated system, a lower initial total acreage equal to $.4 \times 10^6$ acres and a higher initial total acreage of 2×10^6 acres

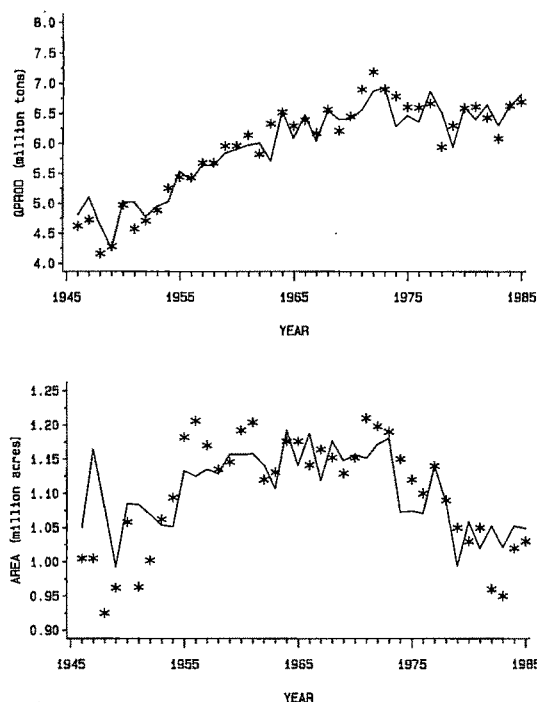


Figure 2. Actual (*) and fitted (—) production (a) and acreage (b) for the California alfalfa market

are also considered. In both cases the system converges to a long-run equilibrium in approximately four to five years after some initial fluctuations about that level.

Quasi-Rational Expectations

This section considers quasi-rational expectations in the sense of Nerlove, Grether, and Carvalho. The authors show (pp. 302–8) that under the rational expectations hypothesis, price expectations can be represented as an ARIMA model with coefficients depending on the coefficients in the structural model. Quasi-rational expectations then consist of estimating ARIMA

Table 2. Price-Elasticities of California Alfalfa Quantity Supplied (naive expectations)

	Alfalfa Price*	PCINDX*	CCINDX*
One-year	.41	-.20	-.36
Long-run	.61	-.29	-.54

Elasticities calculated at 1983–85 average price, yield and acreage levels.

models for the price expectations but ignoring the cross-equation parameter restrictions and then using predicted values from the ARIMA models for expected price.

Quasi-rational expectations are implemented in this study as follows. First, standard time-series methods (Shumway) are used to identify the following models:

log alfalfa price:	ARIMA (2, 2, 0)
log PCINDX:	ARIMA (0, 2, 2)
log CCINDX:	ARIMA (0, 1, 1)

where the entire data series 1945–85 is used in the identification process. Second, for each year t from 1955 to 1985, the above ARIMA models are estimated using data from 1945 to year $t - 1$, and forecasts generated for years s , $s = t, \dots, t + 3$. The four-year average is then used to specify $TRALF_t^*$, $PCINDX_t^*$, and $CCINDX_t^*$ in equations (15) and (16). Finally, the unknown parameters and acreage and removal variables are estimated as before, except that the estimation is carried out for the years 1955–85 instead of 1945–85 under naive expectations.

Table 3 gives the estimated coefficients for the new plantings and removal equations under quasi-rational expectations. The signs of the estimated coefficients in table 3 are roughly comparable to those under naive expectations (table 1). Exceptions are the signs for some age group coefficients in the removal equations and the signs for $CCINDX^*$ in the new plantings equation and the removal equations for age groups 2 and 3. In particular, an increase in the price of competing crops under the quasi-rational expectations hypothesis implies an increase in new plantings area in contrast to what might be expected a priori. The estimated coefficients under quasi-rational expectations are roughly comparable in order of magnitude to the estimated coefficients under naive expectations. Other than that, the most notable differences are higher age-group coefficients and lower $TRALF^*$ coefficients (in absolute value) under quasi-rational expectations.

Table 4 gives the price elasticities for California alfalfa supply under quasi-rational expectations. Here the own-price elasticity is implausibly low (3% in both the short and long run). The production cost elasticities are quite close to those under naive expectations. The elasticity of competing crop prices is positive in contrast to both the results under naive expectations and a priori reasoning. Long-run equilibrium acreage is calculated as 1.04 million acres, which is slightly lower than estimated under naive ex-

Table 3. Estimated New Plantings and Removal Equations for the Alfalfa Supply Response Model (quasi-rational expectations)

Independent Variable	Dependent Variable			
	A_1	R_2	R_3	R_4
A2	-.637 (.107)****	$.58 \times 10^{-1}$ ($.1 \times 10^{-3}$)***	$.39 \times 10^{-2}$ ($.12 \times 10^{-3}$)***	$.35 \times 10^{-2}$ ($.1 \times 10^{-3}$)***
A3	-.738 (.102)***	$.56 \times 10^{-2}$ ($.96 \times 10^{-4}$)***	$.54 \times 10^{-1}$ ($.1 \times 10^{-3}$)***	$.39 \times 10^{-2}$ ($.1 \times 10^{-3}$)***
A4	-.66 (.097)***	$.28 \times 10^{-2}$ ($.9 \times 10^{-4}$)***	$.53 \times 10^{-3}$ ($.1 \times 10^{-3}$)***	.30 ($.98 \times 10^{-4}$)***
Constant	.994 (.086)***	$-.152 \times 10^{-1}$ ($.8 \times 10^{-4}$)***	$-.72 \times 10^{-2}$ ($.98 \times 10^{-4}$)***	$-.53 \times 10^{-2}$ ($.9 \times 10^{-4}$)***
TRALF*	$.40 \times 10^{-4}$ ($.64 \times 10^{-4}$)	$-.34 \times 10^{-5}$ ($.6 \times 10^{-7}$)***	$-.32 \times 10^{-5}$ ($.7 \times 10^{-7}$)***	$-.36 \times 10^{-5}$ ($.6 \times 10^{-7}$)***
PCINDX*	$-.16 \times 10^{-2}$ ($.256 \times 10^{-3}$)***	$.1 \times 10^{-4}$ ($.24 \times 10^{-6}$)***	$.11 \times 10^{-4}$ ($.3 \times 10^{-6}$)***	$.11 \times 10^{-4}$ ($.26 \times 10^{-6}$)***
CCINDX*	$.865 \times 10^{-1}$ (.0453)**	$-.115 \times 10^{-2}$ ($.43 \times 10^{-4}$)***	$-.53 \times 10^{-3}$ ($.5 \times 10^{-4}$)***	$.99 \times 10^{-3}$ ($.46 \times 10^{-4}$)***
R^2	.77	.99	.99	.99
D-W ^b	1.61	2.15	3.1	2.5

Note: Results are from the OLS regressions in the last iteration of the iterative KF/OLS algorithm.

* Standard deviations of the estimated coefficients are in parentheses; single, double, and triple asterisks denote significance at the 10%, 5%, and 1% levels, respectively.

^b D-W is the Durbin-Watson statistic.

Table 4. Price-Elasticities of California Alfalfa Quantity Supplied (quasi-rational expectations)

	Alfalfa Price*	PCINDX*	CCINDX*
One-year	.03	-.25	.09
Long-run	.03	-.29	.11

Note: Elasticities calculated at 1983–85 average price, yield, and acreage levels.

pectations. Forecasted acreage was also estimated for the period 1986–2025 under the same initial conditions and values as before. As with naive expectations, total alfalfa acreage under quasi-rational expectations exhibits a damped cycle with eventual convergence to a long-run equilibrium. However, convergence to long-run equilibrium takes noticeably longer (10–20 years) with supply response estimated under quasi-rational expectations, even assuming constant expected prices and costs during the forecast period.

Conclusions

Previous econometric studies of perennial crop supply response have been based on OLS estimates of single-equation, reduced-form regressions for aggregate output or acreage, or OLS

estimates of individual new planting/removal equations. The first approach does not adequately account for age-distribution impacts on new plantings and removals, while the second approach requires detailed data which are not always available. A state-space approach using the Kalman filter is proposed here. The structural model includes age-distribution effects and separate new planting, removal, and age-group dynamics. Both parameter and age-group acreage estimates are obtained from data on total acreage, production, and prices.

The model is applied to alfalfa production in California. Under naive price and cost expectations, the estimated model provides a reasonable fit to the data. The correct signs were obtained and the price coefficients are generally statistically significant. Estimated long-run equilibrium is close to average acreage over the sample period. The results suggest that alfalfa supply is inelastic in both the short- and long-run. Total acreage converges to long-run equilibrium under constant prices, but with some overshooting. The model was also estimated assuming quasi-rational expectations. The estimated coefficients are generally comparable in sign and order of magnitude to those under naive expectations. However, the own-price supply elasticity is much less than that under naive expectations, and the elasticity with respect to competing crop prices is positive. In general, the results here under naive expectations seem

more consistent with a priori considerations than those under quasi-rational expectations.

Age-structured models arise naturally in many areas of natural resource economics (Wilén). Examples include fisheries, forests, and wildlife herds. In many instances births and/or deaths may depend on the age-structure of the population; thus, accounting for age-structure is necessary for obtaining accurate theoretical and empirical models. Age-structured models also arise in other areas of economics, such as capital theory. In a given industry, for example, existing capital may be quite heterogeneous with respect to age. Investment in new equipment therefore depends, in part, on the age-distribution of the existing capital stock. Economic models which explicitly account for the age distribution of physical and biological capital stocks are relatively scarce. This is due in part to difficulties in analyzing such models theoretically (e.g., Cushing) and in part to difficulties in obtaining age-specific data to conduct empirical analysis. Thus, although the focus of this paper is perennial crops, the method can also be applied to estimate age-structured models in other areas of economics when only age-aggregated data is available.³ Current limitations on the use of this method include lack of generally applicable algorithms for state estimation in nonlinear models and parameter estimation in both linear and nonlinear models.

[Received April 1989; final revision received October 1990.]

References

- Akiyama, T., and P. K. Trivedi. "Vintage Production Approach to Perennial Crop Supply: An Application to Tea in Major Producing Countries." *J. Econometrics* 36(1987):133-61.
- Alston, J. M., J. W. Freebairn, and J. J. Quilkey. "A Model of Supply Response in the Australian Orange Growing Industry." *Aust. J. Agr. Econ.* 24(1980):248-67.
- Arak, M. "The Price-Responsiveness of Sao Paulo Coffee Growers." *Food Res. Inst. Stud.* 8(1968):211-23.
- Bellman, R. E., and M. J. Hartley. "The Tree-Crop Problem." Graduate School of Management Work. Pap. Ser. No. 87-25, University of California, Riverside, June 1987.
- Cushing, J. M. "Nonlinear Matrix Models and Population Dynamics." *Nat. Resour. Model.* 2(1988):539-581.
- Dixon, B. L., and R. E. Howitt. "Resource Production Under Uncertainty: A Stochastic Control Approach to Timber Harvest Scheduling." *Amer. J. Agr. Econ.* 62(1980):499-507.
- . "Uncertainty and the Intertemporal Management of Natural Resources: An Empirical Application to the Stanislaus National Forest." Giannini Foundation Monograph No. 38, University of California, Berkeley, Sep. 1979.
- French, B. C., and R. G. Bressler. "The Lemon Cycle." *J. Farm Econ.* 44(1962):1021-36.
- French, B. C., G. A. King, and D. D. Minami. "Planting and Removal Relationships for Perennial Crops: An Application to Cling Peaches." *Amer. J. Agr. Econ.* 67(1985):215-23.
- Hartley, M. J., M. Nerlove, and R. K. Peters, Jr. "An Analysis of Rubber Supply in Sri Lanka." *Amer. J. Agr. Econ.* 69(1987):755-61.
- Judge, G., W. Griffiths, R. Hill, H. Lütkepohl, and T.-C. Lee. *The Theory and Practice of Econometrics*. New York: John Wiley & Sons, 1985.
- Knapp, K. "Dynamic Equilibrium in Markets for Perennial Crops." *Amer. J. Agr. Econ.* 69(1987):97-105.
- Konyar, K., and K. Knapp. "Market Analysis of Alfalfa Hay: California Case." *Agribus.* 4(1988):271-84.
- Nerlove, M., D. M. Grether, and J. L. Carvalho. *Analysis of Economic Time Series: A Synthesis*. New York: Academic Press, 1979.
- Shumway, R. H. *Applied Statistical Time Series Analysis*. Englewood Cliffs NJ: Prentice-Hall, 1988.
- Wilén, J. E. "Bioeconomics of Renewable Resource Use." *Handbook of Natural Resource and Energy Economics*, ed. A. V. Kneese and J. L. Sweeney, vol. 1, chap. 2. Amsterdam: North-Holland Publishing Co., 1985.
- Wolfram, S. *Mathematica: A System for Doing Mathematics by Computer*. Redwood City CA: Addison-Wesley Publishing Co., 1988.

³ Dixon and Howitt (1979, 1980) use a Kalman filter to update estimated volumes and basal areas by age class in a normative study of forestry management. However, they assume that each of these variables is being sampled in each period. This contrasts with the approach here where there is assumed to be no age-specific data on an annual basis and where annual production and acreage data is used to infer age-specific data.

Measuring the Potential Contribution of Plant Breeding to Crop Yields: Flue-Cured Tobacco, 1954–87

Bruce A. Babcock and William E. Foster

Measurements are made of plant breeders' success in increasing potential flue-cured tobacco yields from 1954 to 1987 in North Carolina. Nongenetic sources of yield increases are measured by the yield levels of a standard variety. New genetic material accounted for between 20% and 35% of yield increases on research station plots in the three tobacco-growing regions. Recent genetic contributions have been small. The slowdown is not attributable to a federal tobacco policy shift in 1965 that slowed the growth in grower yields. There is also no evidence that tobacco breeders increased their output in response to the Plant Variety Protection Act of 1970.

Key words: plant breeding, Plant Variety Protection Act, technical change, tobacco.

This paper measures the potential contribution of new genetic material to changes in the per-acre yields of North Carolina flue-cured tobacco between 1954 and 1987. The methodology is generalizable to crops where research station data exist regarding the cultivation of new plant varieties. For agricultural commodities, genetic sources of technical advance occur as new varieties, or cultivars, are produced by private and public plant breeders. Nongenetic sources of technical advance include new production or cultural practices, such as mechanization, fertilization, pest control, and irrigation.

Yields of new tobacco varieties, obtained from university research station data, are used to construct objective measures of tobacco yield increases due to plant-breeding activities. One important issue motivating this analysis is the contribution of research to observed yields of commercial tobacco farms. Therefore, a critical aspect in measuring genetic contribution is the choice of a research-station yield measure that is realistic in terms of grower behavior. For many crops the highest yielding varieties may not be adopted by growers because of the importance

of quality attributes. This is the case with tobacco because of a negative relationship between yields and beneficial leaf chemistry characteristics. Accordingly, this study bases its measure of genetic contribution on the yield levels of varieties that would maximize grower returns, in contrast to basing the contribution on average yield levels of new varieties, or on maximum yields of new varieties.

The use of research station data limits our ability to understand the ultimate effects of new plant varieties. The foremost limitation is that the adoption decision of farmers cannot be considered. Thus, measures of technical change from research station data can, at best, be interpreted as potential technical change. In addition, because the data are not generated in response to the same economic incentives faced by growers, the response of grower yields to changes in input use cannot be measured.

These limitations, however, are balanced by two benefits. First, because this study offers measures of genetic contribution that are independent of farmer adoption rates, the success of research and development in providing potential technical advances can be tested. This is in contrast to those measures derived from rates of cultivation of hybrid varieties (e.g., Heady and Auer, Guise). In addition, simultaneity, caused by the endogeneity of adoption rates, is avoided.

Besides the general question of measuring genetic and nongenetic contributions to tobacco yields, this paper addresses three specific is-

Bruce A. Babcock is an assistant professor, Department of Economics, Iowa State University. This research was completed while he was an assistant professor at North Carolina State University. William E. Foster is an assistant professor, Department of Agricultural and Resource Economics, North Carolina State University. Senior authorship is not assigned.

The authors thank three anonymous *Journal* referees for their helpful comments.

sues. First, by taking advantage of data from several research stations, the analysis can test how contributions of new technology differ across geographic regions. Previous studies of technical change on corn (Offut, Garcia, and Pinar) and wheat (Feyerherm, Paulsen, and Sebaugh) indicate that regional differences can be significant.

Second, the analysis allows a test of the degree to which research and development of new tobacco varieties was influenced by a significant change in economic incentives facing tobacco growers brought about by a change in federal tobacco policy in 1965. Prior to the policy change, supply was restricted through acreage controls; after the change, supply was restricted by marketing quotas (Grise and Griffin). This change in grower incentives altered actual grower yields (Foster and Babcock). Induced innovation theory (e.g., Binswanger and Ruttan) suggests that technical change is influenced by economic variables at the farmer level, but the relative importance to research of variations in farmer incentives is in question.

Third, the analysis allows a test of the influence on new variety yields resulting from a change in direct incentives facing plant breeders brought about by the Plant Variety Protection Act of 1970 (PVPA). The PVPA increased the ability of plant breeders to appropriate the social gains from new varieties, thus increasing the incentives for breeders, especially private firms, to invest in research and development (Perrin and Foster). PVPA effects on private versus public breeding efforts are difficult to distinguish, but in this analysis the distinction is of limited importance because new varieties typically result from private effort.¹

Modeling Technological Contributions to Yield

The use of different amounts and types of production inputs and new production technologies, and the development of new varieties have both contributed to increased yield levels over time. Disaggregating the contribution of genetic

improvements to yields requires observations of the effects of changing production practices on a standard crop variety.² Yields of standard varieties will change from year to year in response to exogenous shocks, (e.g., weather-related variables), and in response to changes in production practices. The yield of new varieties will change in response to the same exogenous shocks and changes in production practices, as well as to changes in genetic makeup.

Generally, one can represent the yield of the standard variety (Y_s) and new variety (Y_n) as $Y_i = Y_i(t, W)$ ($i = s, n$). The time index t represents the systematic influence on yield of technical innovations, both genetic and nongenetic, and changes in W represent all other exogenous influences on yields. Genetic changes do not affect the yield of the standard variety; therefore, the influence of nongenetic technical change on standard variety yields may be represented by $\partial Y_s / \partial t$. Estimates of $\partial Y_s / \partial t$ also represent the total contributions of nongenetic innovations to increased yields of new varieties. Therefore, the contribution of genetic innovations to increased yields can be obtained by subtracting the nongenetic contribution to the standard variety from the total contribution of technical change to new varieties.³

Regional differences may affect the potential contribution of both nongenetic and genetic technical change to commodity yields. Region-specific production practices reflect differing climatic and soil conditions. Official variety tests conducted at North Carolina tobacco research stations explicitly recognize these differences. While the same varieties are tested across regions, the production techniques used at the research stations are chosen to approximate the respective regions' standard or prevailing practices. In addition, the genetic makeup of some new varieties may be better suited for one geographic area than another. Varietal selection for drought resistance, for example, is likely to increase expected yield levels in drought prone re-

² The term "standard variety" denotes a cultivar that is used by crop scientists to gauge the relative contributions of new tested cultivars over a period of time.

³ There are two components to genetic contribution to increased yields. First, holding all other inputs constant; and, second, allowing an interaction between the new genetic input and changes in other inputs. One cannot separate these two components when data reflect an updating of nongenetic techniques. To the grower deciding on a new variety, and hence to the plant breeder, such a distinction may be irrelevant. Indeed, for a given new variety, one component may be negative; but, so long as the other component is positive and sufficiently large, the variety may be adopted. The important question concerns the total yield improvement from the new genetic input relative to the yield under the old genetic input.

¹ As reported by Bowman et al. (table 1, p. 31) of the 105 new cultivars released between 1954 and 1981, 27 were products of public breeding efforts. The most popular varieties with growers were those developed by private firms. For example, in 1989 only one publicly developed flue-cured variety, NC-60, is recorded by county extension agents as being planted in North Carolina, and this variety accounted for only 3% of plantings (Peedin, Smith, and Yelverton, p. 5).

gions more than in regions with regular rainfall or greater reliance on irrigation. Thus, both $\partial Y_s / \partial t$ and $\partial Y_n / \partial t$ may differ by region.

Selecting the New Variety to Represent Genetic Contribution

For most major commodities, private and public research tend to produce several new commercially available varieties in any given year. Table 1 illustrates the average yearly number of new flue-cured tobacco varieties over a number of periods since 1954. What particular new variety, or combination of varieties, should be used to represent genetic contributions to commodity yields? Previous studies, focusing on aggregate state-level measures use average yields of new varieties in a year (e.g., Bowman et al., Heady and Auer). However, other candidates are available. Not all new varieties will be adopted; *ceteris paribus*, lower-yielding varieties will be ignored in favor of the highest yielding, implying that a measure of new genetic material based on average yields of new varieties will tend to underestimate the contribution. In addition, significant research breakthroughs may be overlooked by using the mean yield in measuring the genetic contribution to yields, particularly when many varieties are released simultaneously.

On the other hand, the highest yielding variety is not necessarily the best representative of genetic contribution. Appearance and the ease of product marketing can be important characteristics. In the case of tobacco, a negative relationship exists between alkaloid concentration (a primary determinant of value per pound—increasing alkaloid concentration increases value) and per-acre yields (Matzinger and Mann). If the highest yielding variety is of low quality,

returning little per unit output, then it is unlikely to be adopted. A more economically justifiable measure of genetic contribution is the return-maximizing yield of the new variety.

The profitability of new crop varieties, however, is often difficult to determine. Different varieties may influence grower costs as well as revenues. Gross revenue data based on the government grades of harvested research station tobacco are collected as part of official variety trials; however, data on possible cost differences are not typically available at a research station. If cost differences between varieties at the grower level are minor relative to revenue differences, then the yield associated with the highest per-acre revenue at the research station is the most profitable.⁴

The importance of the measure one chooses as representing the contribution of genetic material to yield growth varies inversely with the correlation between the different measures. If, for example, each year's highest yielding variety for each region is also the most profitable, then for estimation purposes it does not matter what measure one uses. For flue-cured tobacco the correlation coefficients between the yields of the highest yielding varieties and the yields of the varieties with the highest per-acre revenue average 0.98 for North Carolina. The correlation between the average tobacco yields of new cultivars and the yields of the cultivars with the highest per-acre revenue is not much lower at 0.95. This high correlation exists because of the relatively few number of new varieties released

⁴ Although cost differences between varieties at research stations are negligible, cost differences between varieties for growers, who must make the variety-adoption decision, are not necessarily negligible. Different production practices might favor the adoption of different varieties, and vice versa. However, no evidence of this was found, and the assumption appears justifiable.

Table 1. Summary Statistics of North Carolina Official Variety Trials for Flue-Cured Tobacco

Period	Average Number of New Cultivars	Average Yield of New Cultivars ^a	Average Annual Standard Deviation ^b
1954–59	4.8	1908.5	111.9
1960–66	4.3	2443.7	178.1
1967–72	4.0	2512.1	110.7
1973–77	3.0	2878.8	134.2
1978–82	3.0	2945.0	104.0
1983–87	3.2	2840.7	150.4

Source: Calculated from Department of Crop Science, NCSU.

^a The average yield (pounds per acre) is calculated as the average of all yields of new varieties from all the reporting research stations.

^b The average standard deviation is calculated in two steps. First, the standard deviation of yields is calculated from the reported yields in a given year. Second, these standard deviations are averaged over the years in a given period.

each year (table 1). For other major commodities, the correlations are likely lower. For example, between 1971 and 1979, soybean breeders each year produced an average of forty-nine varieties subject to official variety trials (Perrin, Hunnings, and Ihnen, p. 35).

Constructing the Standard Variety

One practical difficulty with using a standard variety is that research station tests on a single standard variety are rarely carried over all years in a sample. A solution is to use overlapping control varieties over a subset of the sample range. Typically, one standard is recorded simultaneously with another. An initial variety is dropped from experimental trials while a second continues, later to be joined by a third, and so on. Using correlations of standard varieties during overlap periods, predictive equations can be estimated to construct a common, synthetic (i.e., hypothetical) standard variety over the entire period.

Flue-Cured Tobacco Data

In 1954, North Carolina State University began its Official Flue-Cured Tobacco Cultivar Variety Trials (NC OVT). An agreement among growers, seed companies, and manufacturers specifies that only varieties tested in the NC OVT shall be released for commercial sale in North Carolina. The data generated by the trials therefore give a comprehensive documentation of the changes in flue-cured tobacco yields since 1954. The trials are conducted at several tobacco research stations representing the major growing "belts" in the state. The data are reported annually prior to the next tobacco planting (Measured Crop Performance: Tobacco).⁵ Monthly

rainfall and irrigation levels at each station are recorded; the weather variables used in this study are the sum of the two. The averages of research station yields within each belt are the basic units of observation.

The standard varieties grown since 1954 are Hicks from 1954 to 1969, NC 95 from 1961 to the present, and NC 2326 from 1964 to the present. To measure the change in yields from non-genetic innovations over the entire sample range requires addressing the change from Hicks to NC 95.⁶ A functional relationship between the two standard varieties was estimated to predict the yield of NC 95 from 1954 to 1960 if it had been grown. Yield observations of NC 95 at each station were regressed on the yield of Hicks and station-specific rainfall variables from 1961 to 1969. ($R^2 = 0.92$.) Yield differences between belts were accounted for by estimating belt-specific intercept terms. Predicted NC 95 values for the three belts from 1954 to 1960 were obtained by applying the estimated coefficients from the OLS regression to observed yields of Hicks and the rainfall variables in those years. These predicted values for NC 95 were added to the series of observed average NC 95 values from 1961 to 1987 to obtain a standardized variety over the entire sample period for each belt.

Numerous advanced breeding lines and commercially available varieties are tested at each research station every year. New varieties are released for commercial sale after being tested for one year as a commercial variety. The potential contribution of new cultivars to commercial yield increases can be estimated by observing trial yields of new varieties. The number of new varieties released ranges from zero in 1963 and 1968 to eleven in 1954. The results presented in this paper are conditional on the new varieties meeting some minimum acceptable quality standard in the 1950s and early 1960s.⁷ Beginning in 1964 these minimum standards became formalized in the Minimum Stan-

⁵ Data from the Border Belt Tobacco Research Station are available every year since 1954. The Eastern Belt is represented by the Upper Coastal Plain Research Station with data from 1954 to 1987, with the exception of 1969; by the Lower Coastal Plain Research Station from 1967 to 1987, with the exception of 1973, 1983, and 1985; and by the Central Crop Research Station from 1956 to 1965, with the exception of 1961. The Old and Middle Belts are represented by the Upper Piedmont Tobacco Research Station from 1954 to 1987, except for 1984 and 1986; and by the Oxford Tobacco Research Station from 1955 to 1987, except for 1962, 1972, 1975, 1979, and 1982. Missing data are the result of complete or major crop failures caused by severe weather or disease infestations. This study combines the Old and Middle Belts into one region that will be called the Old Belt. The practical reason for doing this is the number of missing observations from the research stations in these two belts. A justification for combining the two is that the two belts have similar soil and climate characteristics.

⁶ NC 95 is chosen as the standard variety over NC 2326 because of more years of overlap with observations of Hicks.

⁷ In measuring the contribution of new genetic material to yield increases, three notable varieties were eliminated from the sample. The three are Coker 139 and Coker 140, released in 1954, and Dixie Bright 244, released in 1955. These varieties (especially Coker 139) had significantly higher yields, were significantly easier to cure and handle at harvest time, but they produced significantly lower levels of nicotine and other alkaloids compared to other varieties. Because the tobacco program in the 1950s paid farmers for all production on their acreage allotment (Grise and Griffin), more than 85% of the crop in 1955 and 1956 was planted to these high-yielding varieties (personal communication with Furney Todd, retired professor of agronomy, North Carolina State University). To discourage the production of low quality tobacco, ASCS withdrew price supports for these varieties.

dards Program, and all varieties presently released for commercial sale in North Carolina meet these quality standards.

Results

This section presents estimates of nongenetic and genetic contributions to yield growth. First, yields (lbs./acre) of the standard and new varieties are regressed against weather variables and time using data from three regional research stations. Differential effects of nongenetic technical change across regions are tested. Second, the contribution of genetic change is estimated by regressing the log of the ratio of new varieties to the standard variety against weather variables and time. Tests are made for the differential effects of genetic technical change across regions. Third, the hypothesis that the genetic contribution to yield growth rates of new varieties slowed because of a change in federal tobacco policy is tested. And finally, the hypothesis that the genetic contribution to yield growth rates of new varieties was altered by the Plant Variety Protection Act is tested.

Regression Results

As described in the first section, the contribution of genetics to yield growth is the additional yield growth of newly released varieties over the yield growth of a standard, in this case NC 95. Two possible measures of the breeder-contributed additional yield growth for use in regression analysis are a simple difference between the new variety yields and NC 95, and the log of the ratio of the two yields. The appropriate measure depends on the functional forms explaining the yields of the standard and new varieties. If yields are linear in the included variables with an additive error term, then the yield difference between NC 95 and the new varieties is an appropriate measure of genetic contribution. If the functional forms for the yield equations have a multiplicative error structure, then the log of the ratio is appropriate. A log functional form was considered more appropriate after diagnostic tests revealed the presence of heteroscedasticity when strict linearity was assumed. Results reported here are for the loglinear case, where the log of yields is regressed against weather variables and the log of the time index.⁸

⁸ The quantitative results for the loglinear case (log-yield against log-time) and the semilog case (log-yield against simple time) are similar.

For each of the three regions, the following equations were used to estimate the nongenetic contribution to growth in tobacco yields:

$$Y_{srt} = k_{sr} + f_r(t) + W_{rt}\beta_{sr} + \epsilon_{srt}$$

$$y_{nrt} = k_{nr} + g_r(t) + W_{rt}\beta_{nr} + \epsilon_{nrt}$$

where y_{srt} and y_{nrt} represent annual yields (in log pounds per acre) of the standard (s) and new varieties (n) in region r (r = Eastern, Old, and Border Belts) and year t ; k_{sr} and k_{nr} are constant terms specific to variety and region; $f_r(t)$ and $g_r(t)$ are quadratic functions of time ($t = 1, 2, 3, \dots$); W_{rt} represents a vector of June, July, and August rainfall (in inches); and ϵ_{srt} and ϵ_{nrt} represent possibly contemporaneously correlated errors.

The results are presented in table 2. The set of equations was estimated as seemingly unrelated because the null hypothesis of no contemporaneous correlation between the error terms of the six equations was rejected at the 1% confidence level. All the coefficients on time are positive and significantly different from zero. In addition, seventeen of the eighteen rainfall coefficients are negative.⁹

Testing for Differential Effects of Technical Change

A differential effect of nongenetic technological change across regions was tested by restricting the estimated coefficients relating time to the yields of NC 95 to be the same across the three regions. The null hypothesis that the effects are equal can be rejected at a 1% confidence level with a calculated F -statistic of 5.83 with 2 and 162 degrees of freedom. Thus, the evidence suggests that nongenetic technological progress has had a differential impact on the three growing regions.

Given the log-linear functional forms, the contribution over time of new genetic material can be obtained by regressing the log of the ratio of the yield of new varieties to the yield of NC 95 on the log of time and rainfall.¹⁰ One equation for each of the three growing belts allows

⁹ Rainfall-squared terms were added to the regressions to account for probable positive, but decreasing, returns from rainfall at low levels of rainfall. The contribution of these terms was negligible, perhaps because of irrigation practices at the research stations in drought years.

¹⁰ The hypothesis of an identical response of NC 95 and the new varieties to the rainfall variables within each of the regions was rejected ($F = 3.17$ with 9 and 162 degrees of freedom). Thus, the rainfall variables are included in the genetic contribution regressions.

Table 2. Estimated Regression Coefficients for the System of Six Yield Equations

Variables ^a	Constant	<i>T</i>	<i>MAY</i>	<i>JUNE</i>	<i>JULY</i>	<i>R</i> ²
Eastern Belt						
New variety	7.725 (71.4) ^b	0.144 (6.20)	-0.0195 (1.68)	-0.0232 (2.17)	-0.00436 (0.68)	0.61
Standard variety	7.734 (89.2)	0.0881 (4.72)	-0.0234 (2.54)	-0.00663 (0.78)	-0.00126 (0.24)	0.46
Old Belt						
New variety	7.443 (73.0)	0.212 (10.9)	-0.0255 (2.21)	-0.0194 (2.10)	-0.00123 (0.13)	0.82
Standard variety	7.585 (87.5)	0.169 (10.4)	-0.0256 (2.60)	-0.0296 (3.73)	-0.00890 (1.15)	0.82
Border Belt						
New variety	7.671 (119.)	0.143 (8.83)	0.00034 (0.28)	-0.0152 (2.91)	-0.00052 (0.10)	0.77
Standard variety	7.713 (129.)	0.105 (7.09)	-0.00210 (1.87)	-0.00641 (1.30)	-0.00703 (1.44)	0.70

^a The dependent variables in the regressions are log yields for the standard variety (NC 95) or the new variety series; *T* is the log of an annual time index that *T* = 1 in 1954; *MAY*, *JUNE*, and *JULY* are inches of rainfall plus inches of applied irrigation water for May, June, and July.

^b The absolute value of the estimated *t*-ratios are given in parentheses.

for differential rates of genetic change across belts. The hypothesis of no contemporaneous correlations among the error terms of the equations was again rejected, and the three equations were run as seemingly unrelated regressions.

Table 3 first reports the regression results assuming regional differences. The estimated annual rate of yield increase from the introduction of new genetic material for a given year, say *t'*, is found by taking the derivative of a genetic contribution equation with respect to time (i.e., $\partial \ln [Y_n/Y_s]/\partial t$) and evaluating at year *t'*. For

example, in the Eastern Belt for 1965, *t'* = 17, and the growth rate equals $\partial \ln [Y_n/Y_s]/\partial t = 0.0556/17 \approx 0.0033$. The estimated average annual rate of yield increase is found by averaging the individual growth rates over the sample period. The resulting rate of yield increases for the period 1954 to 1987 are 0.67% for the Eastern Belt, 0.51% for the Old Belt, and 0.47% for the Border Belt. A test that these growth rates are significantly different across equations can be accomplished by restricting the three time coefficients to be equal. The calculated *F*-sta-

Table 3. Estimated Regression Coefficient Measuring the Contribution of Genetic Material to Yields

Variables ^a	Constant	<i>T</i>	<i>MAY</i>	<i>JUNE</i>	<i>JULY</i>	<i>R</i> ²
Eastern Belt	-0.00621 (0.01) ^b	0.0556 (3.20)	0.00362 (0.42)	-0.0166 (2.10)	-0.00328 (0.69)	0.38
Old Belt	-0.146 (2.25)	0.0425 (3.04)	0.00140 (0.20)	0.0103 (1.83)	0.00763 (1.37)	0.33
Border Belt	-0.0283 (0.55)	0.0390 (2.92)	0.00252 (2.75)	-0.00983 (2.45)	0.00484 (1.21)	0.46
No Regional Yield Growth Differences						
Eastern Belt	0.0206 (0.28)	0.0445 (4.15)	0.00467 (0.55)	-0.0167 (2.13)	-0.00346 (0.73)	0.38
Old Belt	-0.151 (2.47)	0.0445 (4.15)	0.00142 (0.20)	0.0103 (1.86)	-0.00745 (1.36)	0.33
Border Belt	-0.0422 (0.91)	0.0445 (4.15)	0.00259 (2.90)	-0.00974 (2.46)	-0.00465 (1.18)	0.45

^a The dependent variables in the regressions are log yields for the standard variety (NC 95) or the new variety series; *T* is an annual time index that *T* = 1 in 1954; *MAY*, *JUNE*, and *JULY* are inches of rainfall plus inches of applied irrigation water for May, June, and July.

^b The absolute value of the estimated *t*-ratios are given in parentheses.

tistic for testing these restrictions is 0.34, with 2 and 81 degrees of freedom. In contrast to non-genetic change, there is no evidence of a differential impact of genetic research effort across the three tobacco growing belts. Table 3 also reports the regression results imposing these restrictions. The average annual rate of yield increase resulting from new genetic material implied by this restricted regression is approximately 0.54%.

The total yield effect from changes in technology is the sum of the genetic and the non-genetic contributions. The average annual rate of increase of NC 95 can be obtained from the regression results in table 2. The estimated annual growth rate for a given year, say t' , is found by $\partial \ln y_t / \partial t$ evaluated at t' . The average growth rate is found by averaging the individual years' growth rates. The yield of NC 95 grew by average annual rates of 1.07% in the Eastern Belt, 2.05% in the Old Belt, and 1.27% in the Border Belt. The percent of average rates of annual yield increases attributable to new genetic material is found by dividing the estimated average growth rate resulting from new varieties by the sum of growth rates from new varieties and nongenetic gains. The results are 34% in the Eastern Belt

(i.e., $.34 \approx .54 / [.54 + 1.07]$), 21% in the Old Belt, and 30% in the Border Belt. The findings here of differential growth rates within a state are similar to those of Offutt, Garcia, and Pinar, who found that growth rates of corn yields differ across crop reporting districts in Illinois.

Figure 1 illustrates how these growth rates affect the yields of the new variety series and NC 95 in the Eastern Belt. The log of the new variety yield series is the broken line in figure 1. The fitted trend line for this series is taken from the coefficients in table 2. The contribution of changes in production practices to the yields of the new varieties in the Eastern Belt (the lower solid line in fig. 1) is found by subtracting the contribution of genetic material. This contribution was calculated from the restricted regression results in table 3. The contribution of genetics is set to zero in 1954.

Testing for a Plant-Breeder Response to Changed Incentives

The decline in yield growth rates of both the standard variety and the new varieties has several explanations. One is that a "natural" slow-

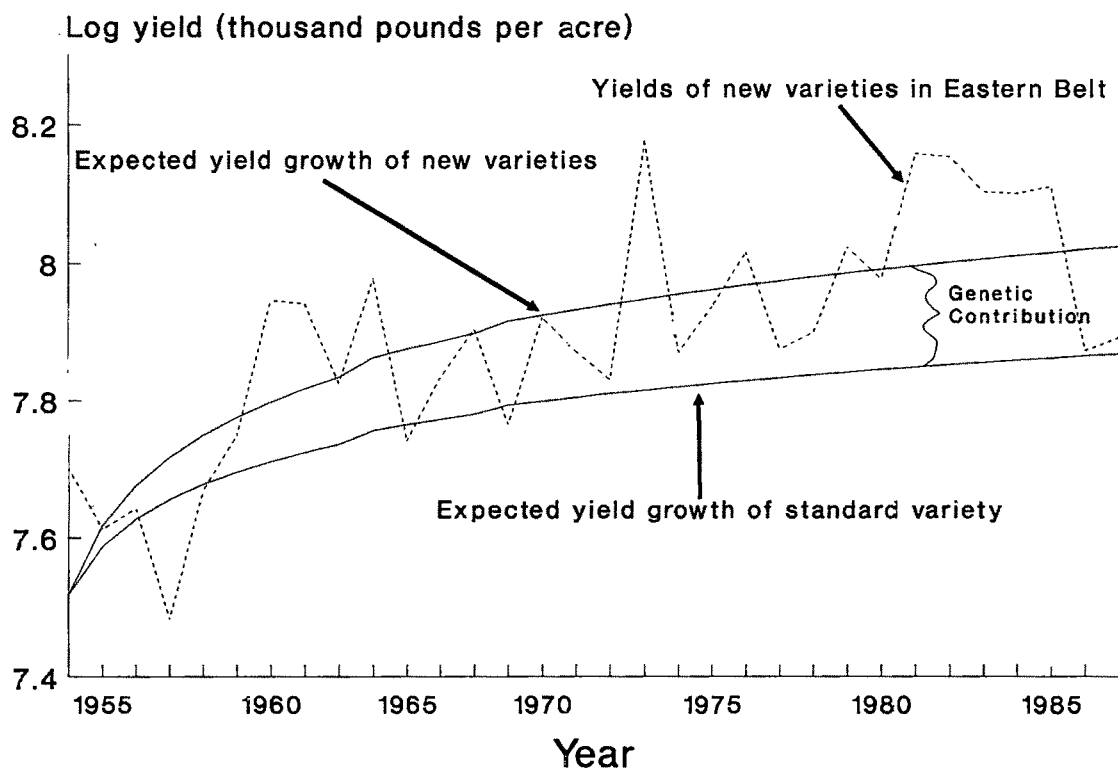


Figure 1. Genetic contribution to yields in the Eastern Belt

down occurred. A slowdown in the growth rate of the yield of NC 95 would appear inevitable given a constant stock of genetic material. A slowdown in the rate of discovery of new genetic material would be expected also once all the possible combinations of existing genetic sources were exhausted.

On the other hand, if plant breeders are responsive to the same economic incentives facing grower clientele groups, then the slowdown might be explained by the major change in flue-cured tobacco policy in 1965. Such a "demand-pull" explanation is consistent with the induced innovation literature. In 1965 acreage allotments were replaced by poundage controls. Foster and Babcock showed that in response to this change in federal policy, expected flue-cured tobacco yields at the county level decreased and the rate of increase of subsequent yields declined. Average county-level tobacco yields and the fitted trend line from their study are shown in figure 2. The average of the three new variety series from the present analysis are also shown in figure 2. A discrete change in the growth rate and/or yield level of new varieties beginning in 1965 would be consistent with the explanation that re-

searchers altered their research goals because of the change in federal tobacco policy.

This hypothesis was tested in the following way. Regime changes for years after 1964 account for possible time lags due to product development.¹¹ Separate tests were performed for shifts in time coefficients and constant terms for periods following each year from 1964 to 1978. For example, to test whether or not a regime change occurred in the development of new varieties six years following the policy change, a test was made for the inclusion of constant and slope dummies that have the value zero prior to 1970.

In the case of an immediate shift in 1965 resultant from the program change, the *F*-statistic testing the null hypothesis of constant intercepts and slopes of the equations measuring the contribution of genetics is 0.39 with 6 and 75 degrees of freedom. Thus, there is no evidence that the change in the tobacco program altered re-

¹¹ Although it typically takes six to seven years to produce genetically stable hybrids, a plant breeder usually has several plant varieties on hand that might serve as candidates for release, so it is not clear how quickly a response by breeders to changed incentives might be observed.

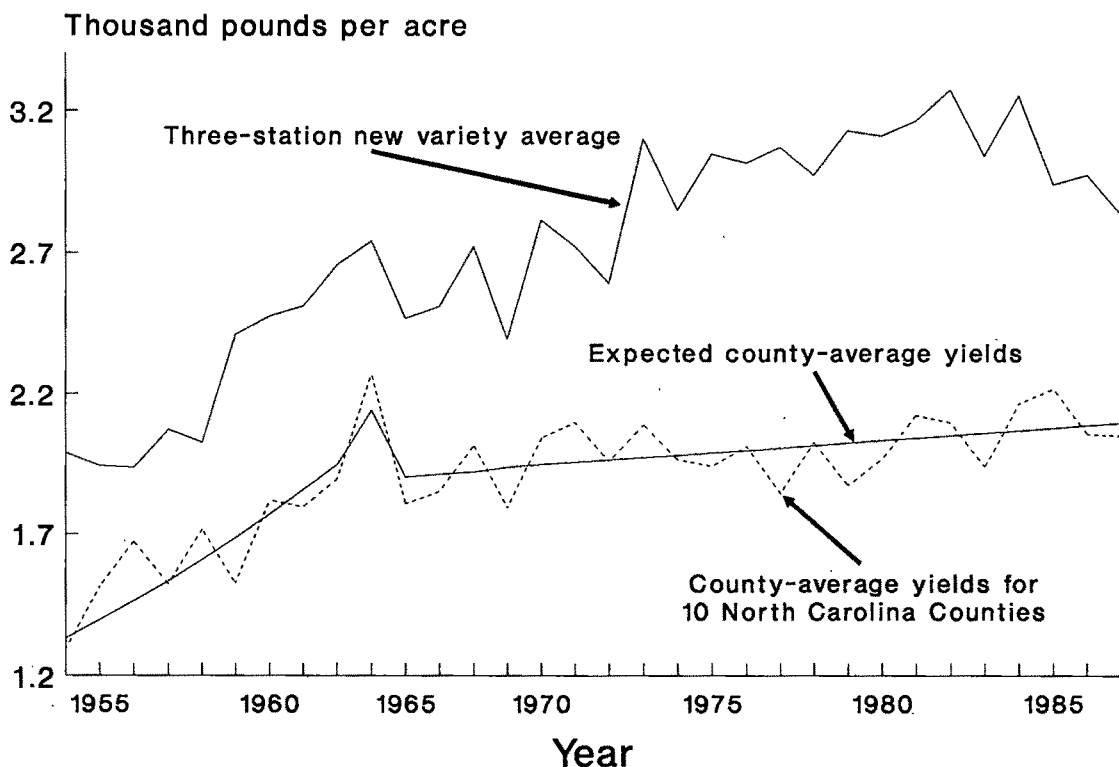


Figure 2. A comparison of county-average yields with research station average yields of new varieties

searchers' delivery of high-yielding varieties beginning in 1965. Indeed, for no year between 1965 and 1978 was the F -statistic testing the null hypothesis of constant intercepts and slopes greater than 1.13.

This negative finding should be interpreted with care. Even though no slowdown in the contribution of genetics could be attributed to the change in the tobacco program, a slowdown in the contribution of yield gains relative to quality gains could occur. The change from acreage allotments to poundage quotas increased the incentives for growers to increase the per pound value of tobacco. Evidence that researchers shifted the mix of yield and quality after the policy change would support the hypothesis that plant breeders do respond to economic incentives felt by their clientele. Such a test is beyond the scope of this study.

The data also allow a test whether the Plant Variety Protection Act (PVPA) of 1970 altered the incentives of tobacco plant breeders similar to the effect on soybean breeders beginning in 1971. No evidence was found that the number of new varieties increased as a result of the PVPA (see table 1). As shown in figure 2, however, the rate of increase of new variety yields in the early 1970s appears to have increased. The basic model presented in table 2 was reestimated eight times allowing for both a shift in the intercepts and a shift in the coefficients on time for eight years after PVPA for the three new variety equations (Perrin, Hunnings, and Ihnen). One year out of eight, 1973, resulted in an F -statistic greater than the 5% critical value of 2.16 with six and 156 degrees of freedom. This evidence is suggestive, but inconclusive, that a response from plant breeders to the adoption of the PVPA may have occurred.

Higher success rates of plant breeding activities after passage of the PVPA would also appear as increases in the genetic contribution to yield increases. As indicated above, however, there is no evidence for a regime change from 1971 to 1978. This finding illustrates the importance of accounting for nongenetic sources of yield variation when testing for changes in the generation of new plant materials. The effect of new cultivars should be estimated as a net effect, not a gross effect.

Conclusion

This paper measures the effect of plant breeding and the development of new varieties on yields

of North Carolina flue-cured tobacco. The main conclusions are as follows: First, genetic innovations have contributed between 20% and 35% of the yield increases on research station plots over the 1954 to 1987 period. Second, recent genetic contributions are small, and the slowdown is not attributable to changes in grower incentives caused by the federal policy shift from land restrictions to marketing restrictions. No evidence suggested that the policy shift, which did have a major effect on grower yields, induced any effective change in new variety research. Third, nongenetic technical change has had differential regional effects, but no regional differences in genetic contribution were found. And last, the establishment of property rights in plant breeding research by the Plant Varieties Protection Act has had no discernable effect on the development of higher yielding flue-cured tobacco varieties.

[Received November 1989; final revision received September 1990.]

References

- Binswanger, H. P., and V. Ruttan. *Induced Innovation*. Baltimore MD: Johns Hopkins University Press, 1978.
- Bowman, D. T., E. A. Wernsman, T. C. Gorbun, and A. G. Tart. "Contribution of Genetics and Production Technology to Long-Term Yield and Quality Gains in Flue-Cured Tobacco." *Tobacco Sci.* 28(1984):30-35.
- Feyerherm, A. M., G. M. Paulsen, and J. L. Sebaugh. "Contribution of Genetic Improvement to Recent Yield Increases in the USA." *Agronomy J.* 76(1984):985-90.
- Foster, W. E., and B. A. Babcock. "The Effects of Government Programs on Flue-Cured Tobacco Yields." *Tobacco Sci.* 34(1990):4-8.
- Grise, V. N., and K. F. Griffin. *The U.S. Tobacco Industry*. Washington DC: U.S. Department of Agriculture, Econ. Res. Serv. Agr. Econ. Rep. No. 589, 1988.
- Guise, J. W. B. "Factors Associated with Variations in the Aggregate Average Yield of New Zealand Wheat." *Amer. J. Agr. Econ.* 51(1969):866-81.
- Heady, E. O., and L. Auer. "Imputation of Production to Technologies." *J. Farm Econ.* 48(1966):309-22.
- Matzinger, D. F., and T. J. Mann. "Genetic Studies on Associations between Nicotine Content and Yield of Flue-Cured Tobacco." Proceedings of the 3rd World Tobacco Scientific Conference, Salisbury, Rhodesia 1963.
- Measured Crop Performance: Tobacco*. Dep. Crop Sci. Res Rep., North Carolina State University, various years.
- Offutt, S. E., P. Garcia, and M. Pinar. "The Distribution of Gains from Technological Advance When Input Quality Varies." *Amer. J. Agr. Econ.* 69(1987):321-27.
- Peedin, G. F., W. D. Smith, and F. H. Yelverton. "Ag-

- ronomic Production Practices." *Tobacco Information*, 1989. North Carolina State University Agr. Extens. Serv., 1989.
- Perrin, R. K., and W. E. Foster. "Economic Measures of the Effect of Intellectual Property Rights on Plant Breeders." Paper presented at the annual meetings of the American Association for the Advancement of Science, New Orleans, 18 Feb. 1990.
- Perrin, R. K., K. A. Hunnings, and L. A. Ihnen. "Some Effects of the U.S. Plant Variety Protection Act of 1970." Dep. Econ. and Bus. Res. Rep. No. 46, North Carolina State University, Aug. 1983.

Using Count Data Models in Travel Cost Analysis with Aggregate Data

Daniel M. Hellerstein

In order to control for censoring and the integer nature of trip demand, the use of count data models in travel cost analysis is attractive. Two such models, the Poisson and negative binomial, are discussed. Robust estimation techniques that loosen potentially stringent distributional assumptions are also reviewed. For illustrative purposes, several count data models are used to estimate a county-level travel cost model using permit data from the Boundary Waters Canoe Area.

Key words: Boundary Waters Canoe Area, count data, negative binomial, Poisson, travel cost.

Estimators of recreational demand models frequently use continuous functional forms, such as ordinary least squares (OLS) on log transformed variables (e.g., Ziemer, Musser, and Hill). However, the nature of trip demand introduces complicating factors. First, trips occur in nonnegative quantities. Failure to control for this censoring will lead to biased estimation. Second, because trips are available only in integer quantities, the usual demand models, which correlate marginal quantity with marginal price, may be inapplicable.

In light of these factors, a natural alternative is to use statistical models that explicitly recognize the "count" nature of trip demand. Several recent papers (e.g., Shaw, Smith, Grogger and Carson, Creel and Loomis) have applied count models to the travel cost model. These works largely have focused on truncated data sets based on choice-based samples. In this study the focus is on the older problem where zero-demanders are included. In particular, the application of several robust estimators of count models to aggregated data will be considered.

The Poisson distribution forms the foundation for the count models examined in this study. Although the Poisson is a convenient distribution

to work with, it imposes some stringent constraints on the demand distribution. In particular, the Poisson distribution assumes the variance of trip demand is equal to the expected value of trip demand. To loosen these constraints, a generalization of the Poisson, the negative binomial, is discussed. Robust estimation procedures, that permit further loosening of a priori assumptions are then reviewed. Permit data from the Boundary Waters Canoe Area are used to examine the effects of these count models on consumer surplus estimates and on coefficient variability.

Theory

In formulating a demand process that yields count data, one must consider that trips are not available in continuous quantities. The integer nature of the data can be explicitly accounted for by modeling the observed number of trips taken (over a season) as the result of many discrete choices (say, one for each day of the season). Under this scenario, count data distributions, such as the Poisson, are an asymptotic outcome.¹

Therefore, in estimating a count model, the analyst is implicitly estimating the "daily" probability of the recreator choosing to visit. Increasing the travel cost will reduce the probability of a visit on any given day. Following Small

Daniel M. Hellerstein is a natural resource economist with the Resources and Technology Division, Economic Research Service, U.S. Department of Agriculture.

Portions of this research were supported by the U.S. Forest Service Rocky Mountain Forest and Range Experiment Station.

The views expressed here are not necessarily those of either the Forest Service or the Economic Research Service.

The author acknowledges Robert Meldelsohn, Michael Moore, George Peterson, and A. Colin Cameron for their advice and assistance.

¹ If the probability of taking a trip on any given day is small, constant and independent of earlier decisions, and if the number of days in a season is large, then the distribution of trips will approach a Poisson process.

and Rosen, integrating over these price changes yields a measure of the compensating variation. Extending these results to the repeated discrete choice context yields a consumer surplus measure over an entire season (Hellerstein and Mendelsohn). The key result is that count data models, as the limit of a repeated discrete choice process, can be used much like continuous models. In particular, integrating under a continuous estimator of predicted demand will yield a measure of consumer surplus.

Given this background, we concentrate on estimating the expected value of trip demand. Furthermore, as a result of repeated discrete choice, the number of observed trips will follow a Poisson distribution. Formally, the expected value of demand is

$$E(Y) = f(P, Z; \beta),$$

where $E(Y)$ is the expected number of trips taken per season, P , Z are explanatory variables including travel cost to site (P) and demand shift variables (Z), such as income and travel costs to substitute sites, and β is a vector of coefficients.

The Poisson probability distribution of demand is

$$(1) \quad \text{Prob}(Y = n; n = 0, 1, 2, \dots) = \exp(-\lambda)\lambda^n/n!$$

with $\lambda = f(P, Z; \beta)$.

The Poisson is a single parameter distribution with expectation and variance both equal to λ (Mood, Graybill, and Boes). Although n is a nonnegative integer, λ must be a strictly positive real number.

The Poisson model is solved by estimating β , say β^* , in $\lambda^* = \lambda(P, Z; \beta^*)$. The estimated value of λ , λ^* , is interpreted as the predicted expected value (and variance) of demand. The predicted expected value of consumer surplus, $E(CS)$, is then computed via the usual integration:

$$(2) \quad E(CS) = \int_{P_{obs}}^{P_{max}} \lambda(P, Z; \beta^*) dP,$$

where P_{obs} is observed price, and P_{max} is a choke price, possibly ∞ .

This paper will focus on this measure of consumer surplus estimated using the predicted mean of demand. Criticisms of this approach in the continuous context (see Bockstael and Strand; or Adamowicz, Fletcher, Graham-Tomasi) stress the importance of considering the source of error. However, if measurement error is small, and

if the measure of expected demand is unbiased, then this measure of consumer surplus will be a good approximation to the expected value of consumer surplus.²

Extending the Poisson: The Negative Binomial

A drawback to the Poisson model is the implied assumption that $E(Y)$ and $\sigma^2(Y)$ are equal. Furthermore, Poisson "regressions" allow no random component in the λ estimator; the $\lambda = \lambda(P, Z; \beta^*)$ relationship does not contain an error component.

The negative binomial count model is often used to relax this unlikely condition of perfect knowledge of the λ estimator and to permit more flexible variance/mean relationships. Following Cameron and Trivedi, the negative binomial is derived as a compound Poisson distribution, where λ is assumed to be distributed as a gamma random variable.³ Integrating over this distribution of λ yields the two parameter negative binomial. Formally,

$$\text{Prob}(Y = n, n = 0, 1, \dots) = \frac{\Gamma(n + \nu)}{\Gamma(n + 1)\Gamma(\nu)} \left(\frac{\nu}{\nu + \mu} \right)^\nu \left(\frac{\mu}{\nu + \mu} \right)^n,$$

with

$$E(Y) = \mu \text{ and } \sigma^2(Y) = \mu + \frac{\mu^2}{\nu}.$$

The variance to mean ratio of the negative binomial is a decreasing function of ν . As ν approaches infinity, the negative binomial collapses to the Poisson; hence the Poisson is nested within the negative binomial.

In terms of the repeated discrete choice framework, the negative binomial admits that

² The concern is that $E(CS(Y))$ may not equal the $CS(E(Y))$, where the latter is what the analyst computes. However,

$$\int_{-\infty}^{\infty} \int_{P_0}^{P_1} d(x; \beta, \epsilon) dP dF(\epsilon) = \int_{P_0}^{P_1} d^*(x; \beta) dP$$

with $d^* = E(Y)$. Therefore, as long as $E(Y)$ is unbiased, then $E(CS(Y))$, the left-hand side of the above equation, will approach $CS(E(Y))$, the right-hand side. It must be stressed that d^* is an unbiased estimate of the expected value of demand. In particular, d^* is assumed to explicitly account for censoring. When this constraint cannot be guaranteed, an alternative form for $d^* = E(Y)$ (such as suggested by Maddala, p. 158) should be used.

³ An alternative derivation can also be found in Hausman, Hall, and Griliches. The Cameron and Trivedi formulation is more flexible, including the Hausman, Hall, and Griliches specification as a special case.

the underlying daily probability of visiting may be randomly distributed. More concisely, each individual is assumed to draw a value for her daily probability at the beginning of the season. Knowledge of the random process generating these daily probabilities is not required, so long as the net result is a gamma distribution of λ , conditional on the exogenous variables.

Estimation

In addition to their appealing statistical properties, the Poisson and negative binomial have several useful empirical properties:

(a) The sum of W independent Poisson variates is also Poisson distributed, with parameter $\lambda_w = \sum_i^W \lambda_i$. Thus, the distribution of visits from the aggregate of W individuals is $\text{prob}(Y = n) = (e^{-\lambda_w})(\lambda_w)^n/n!$. This adding-up property facilitates the use of aggregate data, given knowledge of population size.

(b) If a constant term is included in the function describing λ , the sum (over all observations) of observed demand will equal the sum of predicted demand.

(c) Zero values are admissible. These properties also hold for the negative binomial, with λ replaced by μ .

To insure that λ (or μ in the negative binomial) is strictly positive, it is postulated that

$$(3) \quad \lambda(P, Z; \beta) = \exp(\beta_0 + \beta_P P + \beta_Z Z).$$

These count data models are estimated via maximum likelihood (ML) techniques. The Poisson is readily estimated using the Newton-Raphson technique. The negative binomial, especially its Hessian matrix, is more complicated and is usually solved with a quasi-Newton method, such as the BHHH or the DFP algorithms (Judge et al.).

Maximum likelihood estimation assumes that the postulated distribution is indeed correct. This assumption may impose some stringent requirements, such as the $E(Y) = \sigma^2(Y)$ criteria of the Poisson. The distributional sensitivity of these models raises concerns about robustness. How badly will these models fail if the true probability distribution deviates from the assumed distribution?

The consequences of these assumptions, and possible means of relaxing them, have been studied by a number of authors. For count data models, the work of Gourieroux, Montfort, and

Trognon is especially useful.⁴ They introduce the concept of pseudo- (PML) and quasi-generalized pseudo (QGPML) maximum likelihood estimation. Basically, they show that if functions describing the true mean and true variance of the dependent variable (say, the number of trips) are known [say, equal to $f(p, z; \beta)$ and $g(p, z; \beta, \alpha)$, respectively]; then the Poisson and negative binomial will be consistent, regardless of the underlying distribution. Several interesting results are obtained at the empirical level:

(a) For the Poisson, the estimates of β from ML are consistent, and analytically equivalent to the PML estimates. However, the ML estimate of the covariance matrix of β is too small. The PML estimator can be used to compute a consistent estimate of the covariance (CV) matrix, which turns out to be a function of the "true variance" [the function $g(P, Z; \beta, \alpha)$ introduced above].

(b) For the negative binomial, the ML estimate of β is not consistent if the distributional assumptions do not hold. A two-step QGPML estimator can be used to consistently estimate both α and β . In this case, both β and the CV matrix are functions of the true variance.

An Application to the Boundary Waters Canoe Area

1980 permit data from the Boundary Waters Canoe Area (BWCA) in northern Minnesota will be used to compare the various count models. Each permit contains the ZIP code of the group leader, which can be used to obtain distance-to-site information as well as socioeconomic variables either at the zip code level or at the county level. Because no other information is available, especially information on number of prior trips, the permits must be aggregated. County-level aggregates are used for this analysis. Therefore, the dependent variable is number of trips per county.

A total of 27,433 overnight permits were aggregated into the 1,396 counties within 1,000 road miles of the BWCA.⁵ About half of these

⁴ Other works in this field include White, who offers a general discussion of the consequences of using an incorrect distribution function, and McCullagh and Nelder, who embed count models in the framework of generalized linear models. An interesting summary of count models as weighted, iterative linear estimators is presented by Hall, Hausman, and Griliches.

⁵ All overnight visitors are required to obtain a permit. The entire data base of permits comprises a full census of overnight visitors, who account for about 50% of BWCA use (in RVD terms).

counties have zero visits. Travel cost was based on \$0.076/mile (1980\$) plus a time cost computed using one-third of the per capita wage rate multiplied by average group size (4.0); the result was an average cost per mile of approximately \$0.20.

Table 1 contains the results from several models. An exponential form is used, with an individual's expected demand equal to $\exp(X\beta)$, where X is a vector of exogenous variables. Because aggregate data are used, the population of each county must also be considered. For the count models, the adding-up property suggests use of population as a multiplicative weight. Thus, $E(Y_{\text{county}}) = POP * \exp(X_{\text{county}}\beta)$; Y_{county} is the aggregate number of visits from the county, POP is the population of the county, and X_{county} is a set of county-level exogenous variables (see table 1 for a description of the exogenous variables). For comparative purposes, a simple per capita semilog model is also estimated. Consumer surplus (CS) estimates (in 1980\$) are computed using equation (2), with equation (3) used for λ^* .⁶ In all models, the reported CS estimates use all 1,396 counties.

The most noticeable results are that the own-price coefficient ($BWTC$) is negative and significant for all models. This is especially true for the Poisson model. However, if the Poisson assumption of $E(Y) = \sigma^2(Y)$ is incorrect, then the standard errors generated by the ML estimator of the Poisson will also be incorrect. To test this assumption, a score test devised by Lee is computed. This test is normally distributed under the null hypothesis that the Poisson model is correct. The results of this test, and the large value of the t -statistic for α in the negative binomial model, indicate that the Poisson is inappropriate [that $E(Y)$ does not equal $\sigma^2(Y)$]. Thus, the CV matrix computed by the ML estimator is incorrect, suggesting use of the PML estimator for the Poisson model. In the PML estimator, $BWTC$ is still significantly negative, but several of the demand shifters (the age variables, percent unemployment, and percent poverty) are insignificant at the 95% level.

The negative binomial model returns qualitatively similar coefficients for $BWTC$, % COL -

$LEGE$, and $INCOME$. The sign on the substitute price ($APTC$) is now positive, the theoretically anticipated sign. Also, the age variables are significant.

Within count models, the effect on consumer surplus is on the order of 25%. A much greater change (50%–100%) occurs between count and semilog OLS models. In RVD terms, given an average CS of \$1.5 million, and a total of 25,000 groups of four individuals spending four days in the wilderness, the average RVD value will be about \$4.00 (a fairly small value).

The semilog OLS model uses two heuristics: zero observations are dropped when coefficients are computed, and a bias correction factor is computed. The drop zeros rule ensures computability of the model. Since it discards information about nonvisitors, a potential for bias exists. Alternatively, the semilog model could be estimated using a nonlinear maximum likelihood technique (Creel and Loomis). While such an approach permits zero visits, the treatment of demand shocks resulting from fluctuations in the error term (presumably caused by changes in unobservable factors) is not consistent with demand shocks caused by fluctuations in observed exogenous variables; with changes in unobservable factors having an additive impact, changes in observable variables have a proportional impact. Although this feature may or may not be appealing, for purposes of comparison the simpler drop zeros method is adopted.

The bias correction factor (Stynes, Peterson, and Rosenthal) is a simple multiplier guaranteeing that the sum of observed demand equals the sum of predicted demand. It has substantial impact, leading to a doubling of CS . Alternatively, Bockstael and Strand argue against inclusion of such a bias correction factor. However, the bias correction factor is used since unbiased estimation of $E(CS)$ requires an unbiased estimate of $E(Y)$.

Formal qualitative comparisons between models are presented in table 2 and by examining the η^2 goodness-of-fit statistic in table 1.⁷ Table 2 displays the results of an out-of-sample test created by Ashley and popularized by Shaw.

⁶ To maintain consistency with the definition of the market area (all counties within 1,000 miles), a choke price equal to the maximum price in the sample is used. Alternatively, a choke price of infinity could be used. However, when CS is computed using a choke price of infinity, the results differ by less than 1%. Note that the market area is limited in order to limit the bias resulting from visitors who partake in multiple-destination trips.

⁷ The η^2 statistic can be described as a measure of the correspondence between observed and predicted values. It is related to the familiar R^2 statistic. Specifically, $R^2 = ESS/TSS$ and $\eta^2 = 1 - RSS/TSS$, with ESS the explained sum of squares, TSS the observed (total) sum of squares, and RSS the residual sum of squares. In linear models, these two statistics are analytically equivalent, but they may diverge in nonlinear models. For further discussion of goodness of fit statistics in nonlinear models, see Peterson and Stynes.

Table 1. Results of Count Models, 1980 BWCA Permit Data

	Poisson ML ^a	Poisson PML ^a	Negative Binomial ML ^b	Negative Binomial QGPMML ^b	Semilog OLS (discard zeros) ^c
Constant	-5.14 (-23.0)	-5.14 (-2.45)	-5.42 (-9.3)	-5.39 (-4.2)	-7.62 (-8.9)
BWTC	-0.0208 (-179.0)	-0.0208 (-15.0)	-0.0158 (-79.0)	-0.0158 (-27.0)	-0.0111 (-27.2)
APTC	-0.00219 (-15.0)	-0.00219 (-1.9)	0.00203 (7.4)	0.00209 (4.5)	0.00229 (6.5)
Income	0.000328 (44.0)	0.000328 (5.7)	0.000316 (22.0)	0.000315 (7.9)	0.00017 (4.45)
%Unemploy	1.64 (3.7)	1.64 (0.46)	-4.40 (-4.9)	-4.5 (-2.04)	-0.614 (-0.45)
%Poverty	-2.58 (-7.9)	-2.58 (-1.14)	-0.346 (-0.67)	-0.381 (-0.28)	1.70 (1.69)
%College	5.51 (20.0)	5.51 (2.3)	2.6 (3.7)	2.60 (1.57)	4.09 (3.3)
% < 17	-1.06 (-2.5)	-1.06 (-0.26)	-5.46 (-1.2)	-3.55 (-1.27)	1.49 (0.76)
% > 65	-0.0256 (-0.054)	-0.0256 (-0.006)	-5.47 (-4.2)	-5.58 (-2.05)	1.62 (0.87)
α^d		9943	7915 (20.0)	29855	
Lee test ^e	292.0	292.0	1,680	1,689	2,538
CS (k\$)	1,317	1,317	0.67	0.66	0.68
η^2	0.92	0.92			

Notes: Dependent variable: count models, number of visits per county; semilog OLS model, per capita visitation. Functional form (W = population): for count models, $E(Y) = W \cdot \exp(X\beta)$; for semilog OLS, $E(\ln(Y/W)) = X\beta$. Here, 27,433 permits are aggregated by county into 1,396 observations. Of these 1,396 observations, 647 observations contain zero visits (zero permits). T -statistics are displayed in parenthesis. For Poisson and QGPMML count models, the standard errors are computed using the inverse Hessian. The negative binomial ML uses $\ln(Z'/Z)$, with $Z = \partial L / \partial \beta$. Independent variables: BWTC is travel cost to BWCA; APTC, travel cost to Algonquin Provincial Park, a substitute site in southern Ontario; INCOME, per capita income; %UNEMPLOY, % of labor force currently unemployed; %COLLEGE, % of population older than 16 with college degree; %POVERTY, % of households with income below poverty level; % < 17, % of population less than 17 years old; and % > 65, % of population greater than 65 years old. The GAUSS programs used to estimate these models are available from the author.

^a The Poisson ML and the Poisson PML models yield analytically equivalent estimates of β .

^b The NEGBIN II model is used, with the variance to mean ratio linear-in-the-mean: $\sigma^2(y) = E(y)[1 + \alpha E(y)]$.

^c A bias correction factor of 0.74, computed to force the sum of predicted visits to equal the sum of observed visits is incorporated into the constant.

^d For the PML Poisson and the QGPMML Negative Binomial, α is computed in separate regression; it is not estimated using the likelihood function, hence standard errors are not available. Note that in the MLE negative binomial, α insignificantly different from 0 implies that the variance equals the mean.

^e Under the null hypothesis the Poisson is the correct distribution, implying that the mean equals the variance, the Lee test (Lee) is distributed as a standard normal. A one-tailed test is appropriate, with critical value at 1.64.

Table 2. Ashley Test of Model Quality

Model 1 ^a	Model 2	The superior model is . . .
Poisson	NegBin	inconclusive
Poisson	QGPML	inconclusive
Poisson	Semilog+	Poisson
NegBin	QGPML	NegBin
NegBin	Semilog+	NegBin
QGPML	Semilog+	QGPML

Note: The Ashley test (Ashley, or Shaw) uses out-of-sample data (1981 data) to directly compare the predictive power of two different models.

^a Models: Poisson, Poisson model; NegBin, negative binomial estimated using ML; QGPML, negative binomial estimated using QGPML; Semilog+, OLS on semilog model with a "bias correction" factor.

This test is essentially a nonparametric comparison of the goodness of fit of two different models, using out of sample data. In this case, 1981 data are used as the out-of-sample data, and 1980 data are used to estimate the coefficients.⁸

The η^2 goodness-of-fit statistics suggest that fit is fairly good, especially for the Poisson model. The negative binomial have predictive accuracy similar to the bias-corrected semilog. The results of the Ashley test, although not conclusive, suggest that count models are superior to the bias-corrected semilog model. Within count models, the evidence is weaker. For example, the ML negative binomial is considered superior to the QGPML negative binomial. It is interesting that a Hausman specification test (Hausman) comparing the QGPML and the ML estimators fails to reject the consistency of the ML estimator; the result indicates that the negative binomial distribution is correct.

These results indicate that count models outperform the drop-zeros semilog model. Within count models, the Poisson outperforms the negative binomial in predictive accuracy but produces incorrect measures of variability, throwing the Poisson t -statistics into doubt. The PML estimator of the Poisson can be used to address this failure while maintaining the predictive power of the Poisson. The QGPML negative binomial is similar to the ML negative binomial, suggesting either that the ML estimator be used (on

efficiency grounds) or that the QGPML estimator be used (on ease of computation grounds).

Concluding Comments

The intrinsic nature of site visitation suggests that demand models based on continuous functional forms are inappropriate because they fail to recognize the count nature of trip making. To account for this feature, two count models, the Poisson and negative binomial, are reviewed. At an empirical level, besides matching the integer quality of trip demand, these count distributions explicitly account for censoring at zero. Furthermore, a behavioral justification for their use, based on a repeated discrete choice process generating trip demand, can be derived.

An application to permit data from the Boundary Waters Canoe Area reveals that the choice of estimator can have substantial impact, especially on consumer surplus estimates. In particular, the drop-zero semilog OLS yielded estimates approximately 50% larger than the Poisson. The consumer surplus differences among the several count models were smaller. This suggests that for welfare purposes, the ML Poisson may be adequate. However, coefficient values did vary, especially for demand shift variables. More important, t -statistics for all variables were quite different across count models, with the robust estimators (such as the PML Poisson) generally returning much smaller t -statistics. These results suggest caution when interpreting coefficient values from maximum likelihood estimators.⁹

Although the application of count models to travel cost analysis is becoming increasingly popular, the existence of robust estimators for both the Poisson and negative binomial has not been exploited. These robust estimators reduce the extent of a priori knowledge required for consistent estimation of the coefficient vector and its covariance matrix.

A discussion of the aggregation issue also is in order. First, consider the robust estimators used here. They all require that the functional form describing expected demand is correct. In the context of aggregate data, this requires that the

⁸ The Ashley test uses out-of-sample data. To test that 1981 data is out of sample but not generated by a different model, parameter stability is tested. Specifically, a Chow test of parameter equality, using 1980 and 1981 data with a Poisson PML model, failed to reject the null of parameter stability; with an $F(9, 2774)$ -statistic of 0.92, well below the 95% cutoff level at 1.88. However, the MLE model did reject the null, with an F -statistic of 42.0.

⁹ To the extent that variance of consumer surplus is a function of the covariance of the coefficient vector, this result also suggests caution when qualitative comparisons are formed. See Bockstael and Strand or Adamowicz, Fletcher, and Graham-Tomasi for further discussion of the effects of uncertainty on the variance of consumer surplus estimates.

X_{county} measures be representative, in the sense of Deaton and Muelbauer (p. 149), of the county. If they are not, then $W^* \exp(X_{\text{county}}\beta^*)$ will not equal $\sum_{i=1}^W \exp(x_i\beta^*)$, and the requirements for robust estimation will be violated.

This weakness of aggregate data sets must be measured against the weaknesses of alternative methods, such as the sample selection models of Shaw or the truncated count models of Creel and Loomis. These models admit the influence of nonvisitors in a limited fashion. First, for tobit-like estimators and for Poisson-based models, incorrect specification of higher moments will bias estimates of demand parameters. Although the negative binomial is robust to misspecification of higher moments (Grogger and Carson), all these models are sensitive to the presumption that nonvisitors possess the same demand parameters as visitors. To the extent that this is not true, truncated models may be more biased than aggregated models. In other words, aggregate models permit nonvisitors to influence estimation, so that the resulting parameters are a reduced form incorporating information on both visitors and nonvisitors. For many purposes, such as calculating the CS for a new population, such parameters may be superior to those produced by truncated models.¹⁰ In short, aggregate analysis is not necessarily dominated by site-based samples estimated with econometric techniques that recognize truncation.

[Received April 1990; final revision received August 1990.]

References

- Adamowicz, Wiktor, Jerald Fletcher, and Theodore Graham-Tomasi. "Functional Form and the Statistical Properties of Welfare Measures." *Amer. J. Agr. Econ.* 69(1989):414-21.
- Ashley, R., C. W. J. Granger, and R. Schmalensee. "Advertising and Aggregate Consumption: An Analysis of Causality." *Econometrica* 48(1980):997-1016.
- Bockstael, Nancy, and Ivar Strand. "The Effect of Common Sources of Regression Error on Benefit Estimates." *Land Econ.* 63(1987):11-20.
- Cameron, Colin, and Pravin Trivedi. "Econometric Models Based on Count Data: Comparisons and Applications of Some Estimators and Tests." *J. Appl. Econometrics* 1(1986):29-53.
- Creel, Michael, and John Loomis. "Theoretical and Empirical Advantages of Truncated Count Data Estimators for Analysis of Deer Hunting in California." *Amer. J. Agr. Econ.* 72(1990):434-45.
- Deaton, A., and J. Muelbauer. *Economics and Consumer Behavior*. New York: Cambridge University Press, 1980.
- Gourieroux, C., A. Monfort, and A. Trognon. "Pseudo Maximum Likelihood Methods: Applications." *Econometrica* 52 (1984):701-20.
- Grogger, J. T., and R. T. Carson. "Models for Counts from Choice Based Samples." Dep. Econ. work. pap., University of California, July 1988.
- Hall, Bronwyn Jerry Hausman, and Zvi Griliches. "Patents and E&D: Is There a Lag?" *Int. Econ. Rev.* 27(1986):255-82.
- Hausman, Jerry. "Specification Tests on Econometrics." *Econometrica* 46(1978):1251-70.
- Hausman, Jerry, Bronwyn Hall, and Zvi Griliches. "Econometric Models for Count Data with an Application to the R&D Relationship." *Econometrica* 52(1984):909-37.
- Hellerstein, Daniel, and Robert Mendelsohn. "Modeling Recreational Demand as a Repeated Discrete Choice Process: The Case for Count Models." Washington DC: U.S. Department of Agriculture, Econ. Res. Serv. work. pap.
- Judge, G., W. Griffiths, R. Hill, H. Lütkepohl, and T. Lee. *The Theory and Practice of Econometrics*. New York: John Wiley & Sons, 1980.
- Lee, Lung-Fei. "Specification Test for Poisson Regression Models." *Int. Econ. Rev.* 27 (1986):689-706.
- Maddala, G. S. *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge: Cambridge University Press, 1983.
- Mood, Alexander, Franklin Graybill, and Duane Boes. *Introduction to the Theory of Statistics*. New York: McGraw-Hill Publishing Co., 1974.
- Mullahy, John. "Specification and Testing of Some Modified Count Data Models." *J. Econometrics* 33(1986):341-65.
- McCullagh, P., and J. A. Nelder. *Generalised Linear Models*. London: Chapman and Hill, 1983.
- Peterson, George, and Daniel Stynes. "Evaluating Goodness of Fit in Nonlinear Recreation Demand Models." *Leisure Sci.* 8(1986):131-47.
- Shaw, Dai Gee. "On-Site Sample's Regression: Problems of Non-Negative Integers, Truncation, and Endogenous Selection." *J. Econometrics* 37(1988):211-23.
- Small, Kenneth, and Harvey Rosen. "Applied Welfare Economics with Discrete Choice Models." *Econometrica* 49(1981):105-30.
- Smith, V. Kerry. "Selection and Recreation Demand." *Amer. J. Agr. Econ.* 70(1988):29-36.
- Stynes, Daniel, George Peterson, and Donald Rosenthal. "Log Transformation Bias in Estimating Travel Cost Models." *Land Econ.* 62(1986):94-103.
- Ziemer, Rod, Wesley Musser, and R. Carter Hill. "Recreation Demand Equations: Functional Form and Consumer Surplus." *Amer. J. Agr. Econ.* 62(1980):136-41.

¹⁰ Furthermore, extensions to the Poisson that permit systematic differences between visitors and nonvisitors, such as proposed by Mullahy, are estimable with nontruncated samples (although the robustness of these models requires further investigation).

Errata

Two items were inadvertently not included in the 1990 "Report of the AAEA Awards Committee" published in the December 1990 issue of the *AJAE*. One of these is presented as table 1, below. It is the listing of the number of submissions for each award, the number of awards,

and the number of individuals receiving recognition.

The second item is a listing of the individuals who served on the association's awards committees that year. This list was published in the *AAEA Newsletter*, September-October 1989.

We apologize for any inconvenience these omissions may have caused.

Table 1. Numbers of Nominations and Winners for AAEA Awards, 1989 and 1990

Award	Number of Nominations		Number of Winners	
	1989	1990	1989	1990
Distinguished Undergraduate Teaching				
<10 years	5	9	1	1
≥10 years	5	8	1	1
Distinguished Extension Program				
Individual	5	6	1	1
Group	6	4	1	1
			(8) ^a	(3) ^a
Distinguished Policy Contribution	5	7	1	1
			(6) ^a	
Publication of Enduring Quality	8	12	1	1
			(2) ^a	(3) ^a
Outstanding <i>Journal</i> Article	Fixed by number of articles in volume		1	1
Quality of Communication			2	2
	30	9	(21) ^a	(5) ^a
Quality of Research Discovery	17	5	1	2
			(4) ^a	(4) ^a
Outstanding Ph.D. thesis	22	28	3	3
Outstanding master's thesis	29	34	3	3

^a Denotes the number of individuals involved in the award.

A Model of Production with Supply Management for the Canadian Agricultural Sector: Comment

Carl H. Nelson

In a recent paper Moschini (1988) claims to prove that cost complementarities between outputs do not exist when outputs are joint because of fixed allocatable factors which are normal inputs. This comment demonstrates that Moschini's proposition is incorrect, shows where the proof of the proposition is faulty, and identifies the additional condition that must hold in order for Moschini's proposition to be true. Correct derivation of this proposition is important because technology with fixed allocatable factors and multiple outputs is common in agriculture and standard analyses of multiple output cost functions, such as Sakai's or Lau's, do not apply directly to this technology.

Consider the multiple output cost function used in Moschini (1988) to characterize production which is joint in outputs because of constraints on allocatable fixed factors. Jointness is not present in the primal representation of this technology. The production possibility set can be represented with individual production functions for each output. But outputs are joint in the dual cost and profit functions because they must compete for a constrained quantity of some allocatable input. The joint cost function is defined by Moschini's equation (15), which is reproduced here as equation (1):

$$(1) \quad C(y, w, z) = \min_{(z^i)} \left\{ \sum_i C^i(y_i, w, z^i) \mid \sum_i z^i \leq z \right\},$$

where $y = (y_1, y_2, \dots, y_n)$ is the vector of outputs, w is the vector of prices for variable inputs, z is a scalar representing the total amount of the fixed input that can be allocated to production of the various outputs, and z^i , $i = 1, 2, \dots, n$ is the amount of the fixed input allocated to output i . The joint cost function $C(y, w, z)$ is assumed to be twice continuously differentiable in its arguments. This implies that each of the individual cost functions is twice continuously differentiable in its arguments. And it is assumed that the solution to (1) occurs in the interior of the feasible region.

Carl H. Nelson is an assistant professor, Department of Agricultural Economics, University of Illinois.

This research is part of Project No. 05-334 of the Illinois Agricultural Experiment Station.

The helpful comments of David S. Bullock and three anonymous referees are gratefully acknowledged.

Moschini claims to prove that, given these assumptions and the additional assumption of normal fixed inputs (i.e., $\partial^2 C^i / \partial y_i \partial z^i \leq 0$), cost complementarities between outputs do not exist (i.e., $\partial^2 C / \partial y_i \partial y_j \geq 0$). The proof of this claim appears in Moschini's equations (17) through (19). The error in the proof occurs in equation (18). Moschini's equation (18) is a standard primal comparative static result obtained by totally differentiating the first-order conditions from the cost minimization problem defined in equation (1). A correct derivation of this comparative static result follows.

The first-order conditions for the problem in equation (1) are $\partial C^i / \partial z^i = \theta$ $i = 1, 2, \dots, n$ and $\sum_i z^i = z$, where $\theta = \partial C / \partial z$ is the negative of the quasi-rent of the fixed input z . Total differentiation of these equations with respect to the choice variables, θ , z^1 , z^2 , \dots , z^n , and the exogenous variables z , y_1 , y_2 , \dots , y_n yields

$$(2) \quad \begin{bmatrix} 0 & -1 & -1 & \cdot & \cdot & \cdot & -1 \\ -1 & C_{zz}^1 & 0 & 0 & \cdot & \cdot & 0 \\ \cdot & 0 & C_{zz}^2 & 0 & \cdot & \cdot & 0 \\ \cdot & \cdot & 0 & \cdot & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ -1 & 0 & \cdot & \cdot & \cdot & \cdot & C_{zz}^n \end{bmatrix} \begin{bmatrix} d\theta \\ dz^1 \\ dz^2 \\ \cdot \\ \cdot \\ dz^n \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & -C_{zy}^1 & 0 & \cdot & \cdot & \cdot & 0 \\ \cdot & 0 & -C_{zy}^2 & 0 & \cdot & \cdot & 0 \\ \cdot & \cdot & 0 & \cdot & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & \cdot & 0 & -C_{zy}^n \end{bmatrix} \begin{bmatrix} dz \\ dy_1 \\ dy_2 \\ \cdot \\ \cdot \\ dy_n \end{bmatrix},$$

where $C_{zz}^i = \partial^2 C^i / \partial z^i \partial z^i$, and $C_{zy}^i = \partial^2 C^i / \partial y_i \partial z^i$. The derivative defined in Moschini's equation (18), $\partial z^i / \partial y_m$, can be obtained by premultiplying both sides of equation (2) by the inverse of the matrix on the left-half side of the equation. As is standard in such comparative static derivations the matrix on the left-hand side of the equation is the bordered Hessian of the objective function. The second-order condition of the optimization problem in (1) requires that this matrix be positive definite subject to one constraint, which implies that all border-preserving principal minors of the bordered Hessian must be negative.

Because the matrix on the right-hand side of equation (2) is diagonal the solution for $\partial z^i/\partial y_m$ will only involve the term in the i th row and m th column of the inverse of the bordered Hessian. Letting A represent the diagonal matrix in the lower right-hand corner of the bordered Hessian and ι a column vector of 1's, partitioned inversion allows the inverse of the bordered Hessian to be written as: $[A^{-1} - A^{-1}\iota\iota'A^{-1}/(\iota'A^{-1}\iota)]$, yielding the result that the (i, m) th element of the inverse of the bordered Hessian is

$$(3) \quad \frac{1}{C_{zz}^i C_{zz}^m \left(\sum_{k=1}^n \frac{1}{C_{zz}^k} \right)}$$

for i and m between 2 and n . This means that the correct expression for $\partial z^i/\partial y_m$, instead of Moschini's equation (18), is

$$(4) \quad \partial z^i/\partial y_m = \frac{(\partial^2 C^m/\partial z^{m*}\partial y_m)}{\partial^2 C^i/\partial z^i\partial z^i\partial^2 C^m/\partial z^{m*}\partial z^{m*} \left(\sum_{k=1}^n \frac{1}{\partial^2 C^k/\partial z^{k*}\partial z^{k*}} \right)}.$$

The error in Moschini's argument is the claim that second-order conditions and input normality require the expression in (4) to be negative. While it is true that the assumption of input normality requires the numerator of (4) to be negative, it is not true that second-order conditions require the denominator of (4) to be positive. The second-order conditions allow one of the second derivatives, $\partial^2 C^j/\partial z^j\partial z^j$, to be negative as long as it is smaller in absolute value than all other such second derivatives. This means there are three conditions under which the denominator of (4) can be negative. First, $\sum_{k=1}^n 1/(\partial^2 C^k/\partial z^{k*}\partial z^{k*})$ can be negative if $1/(\partial^2 C^j/\partial z^j\partial z^j)$ for some j not equal to i or m is negative and larger in absolute value than $\sum_{k \neq j}^n 1/(\partial^2 C^k/\partial z^{k*}\partial z^{k*})$. Second, $\sum_{k=1}^n 1/(\partial^2 C^k/\partial z^{k*}\partial z^{k*})$ can be negative if $1/(\partial^2 C^i/\partial z^i\partial z^i)$ is negative and smaller in absolute value than $\sum_{k \neq i}^n 1/(\partial^2 C^k/\partial z^{k*}\partial z^{k*})$. Or third, $\sum_{k=1}^n 1/(\partial^2 C^k/\partial z^{k*}\partial z^{k*})$ can be negative if $1/(\partial^2 C^m/\partial z^{m*}\partial z^{m*})$ is negative and smaller in absolute value than $\sum_{k \neq m}^n 1/(\partial^2 C^k/\partial z^{k*}\partial z^{k*})$. Contrary to Moschini's claim, each of these conditions is allowed by the second-order conditions for the problem defined in equation (1).

The expression in equation (4) is negative, and Moschini's proposition about cost complementarity is true if an additional restriction is imposed upon Moschini's model. The additional restriction is that each of the single output cost functions be convex in z . This assumption requires all of the second derivatives, $\partial^2 C^j/\partial z^j\partial z^j$, to be positive. Assuming convexity of all the single-output cost functions is the same as assuming that all the single-output production functions are concave in z . This assumption implies that the production possibility set is a convex set, which is the sufficient condition for a profit-maximizing choice of fixed and variable inputs to be a long-run equilibrium. Thus, Moschini's proposition

holds when the sufficient conditions for long-run equilibrium hold.

But, jointness caused by fixed allocatable factors is likely to be of greatest interest when fixed assets are not in long-run equilibrium. And there is reason to believe that some fixed assets in agriculture are in disequilibrium. Nelson, Braden, and Roh find evidence that aggregate agricultural machinery and tractors are not in long-run equilibrium. When an asset is not in long-run equilibrium, the only theoretical restriction on $\partial^2 C^i/\partial z^i\partial z^j$ comes from the second-order conditions for an interior solution to (1). And these second-order conditions do not require $\partial^2 C^i/\partial z^i\partial z^j \geq 0$ for all i . That is, they do not require that all of the individual cost functions be convex in z^j . Rather, they rule out "too much" concavity in z^j . In other words, if the cost function for one of the outputs decreases

too rapidly in z , or more than one single output cost function is concave in z , then the first-order conditions for an interior solution do not describe a cost-minimizing fixed input allocation. But, as long as none of the individual cost functions decreases too rapidly in z , an interior solution with joint production is consistent with concavity of one of the single output cost functions in z .

Insight into the restrictions on concavity that are imposed by the second-order conditions can be gained from the following simple three-output example. If $\partial^2 C^1/\partial z^1\partial z^1 = 2$, $\partial^2 C^2/\partial z^2\partial z^2 = -4$, and $\partial^2 C^3/\partial z^3\partial z^3 = 6$, the second-order conditions for cost minimization are violated; but if $\partial^2 C^1/\partial z^1\partial z^1 = 2$, $\partial^2 C^2/\partial z^2\partial z^2 = -1$, and $\partial^2 C^3/\partial z^3\partial z^3 = 6$ the second-order conditions are satisfied and $\partial^2 C/\partial y_1\partial y_3 \leq 0$. Intuitively, this derivative is negative because an exogenous increase in y_3 causes an increase in z^3 and a larger-than-offsetting decrease in z^2 , because of the decrease in the quasi-rent, $q(y, w, z)$, of z . An exogenous increase in the amount of y_3 decreases the quasi-rent of z because an increase in y_3 reallocates z to less valuable uses. The additional z freed by the reduction in z^2 is used to increase z^1 , which reduces $\partial C/\partial y_1$.

If z were in long-run equilibrium, the quasi-rent of z would equal the market price of z , r , and $\partial C/\partial z^j$ would equal $-r$ for all i . Further, the second-order condition for cost minimization would require that $\partial^2 C^i/\partial z^i\partial z^j \geq 0 \forall i$. The fact that the cost of producing a given output can be concave in a fixed input, while it must be convex in variable inputs has bearing on the argument in Moschini (1989) claiming to demonstrate that $\partial^2 C(y, w, r)/\partial y_i\partial y_j \geq 0 \forall i \neq j$. The Moschini (1989) argument for the absence of cost complementarities is critically dependent on the assumption that the production possibility set is convex in the fixed inputs, z . As with Moschini (1988), the

Moschini (1989) argument is not valid if the sufficient conditions for long-run equilibrium, a convex production possibility set, are not satisfied.

The dependence of the Moschini (1989) argument on this assumption highlights the fact that Lau's Hessian identity results must be applied with care to agricultural production. Lau's claims about properties of the Hessians, in particular the claim that the restricted profit function is concave in z , depend upon the assumption that the production possibility set is convex in z . When fixed assets are not in long-run equilibrium convexity may be violated, as it is if one of the individual production functions is not concave in z . Under such conditions, neither Sakai's normal case nor Moschini's proposition about cost complementarity correctly describes the cost and profit interactions between multiple outputs. The derivatives, $\partial^2 C^i / \partial z^j \partial z^l$, are important parameters for technologies joint in outputs because of fixed allocatable factors. The signs of these parameters should be determined

empirically in order to understand the nature of technology joint from fixed allocatable factors.

[Received March 1990; final revision received September 1990.]

References

- Lau, L. "A Characterization of the Normalized Restricted Profit Function." *J. Econ. Theory* 12(1976):131-63.
- Moschini, G. "A Model of Production with Supply Management for the Canadian Agricultural Sector." *Amer. J. Agr. Econ.* 70(1988):318-29.
- . "Normal Inputs and Joint Production with Allocatable Fixed Factors." *Amer. J. Agr. Econ.* 71(1989):1021-24.
- Nelson, C., J. Braden, and J. Roh. "Asset Fixity and Investment Asymmetry in Agriculture." *Amer. J. Agr. Econ.* 71(1989):970-79.
- Sakai, Y. "Substitution and Expansion Effects in Production Theory." *J. Econ. Theory* 9(1974):255-74.

A Model of Production with Supply Management for the Canadian Agricultural Sector: Reply

Giancarlo Moschini

Nelson correctly points out an error in equation (18) of Moschini (1988). The formula that I gave is valid for the two-output case; but, for the general n -output case, the correct expression is given by Nelson's equation (4). I should be thankful for his careful reading of my paper. Regrettably, Nelson goes on to make a number of erroneous inferences which lead to unfounded criticisms of my analysis. I will show that the result he questions is valid without any stronger conditions than those contained in my paper. In particular, Nelson confused two basic properties of production technologies. He claims that my result requires concavity of the production function when, in fact, it is valid as long as the production function is quasi-concave. Below, I will prove that quasi-concavity suffices, and I will show that violation of quasiconcavity leads to inconsistent definitions of input normality. Finally, I will show that Nelson's analysis of the relationship of my results to long-run equilibrium conditions is incorrect.

The context is a multiproduct firm with an allocatable fixed input z , n production lines with individual production functions $y_i = f^i(x^i, z^i)$, where (x^i, z^i) are inputs allocated exclusively to the i th production line, and where one of the inputs must satisfy the constraint $\sum_i z^i \leq z$. In this setting, under weak regularity conditions and given a vector of variable input prices w , one can define individual cost functions $C^i(y_i, w, z^i)$ and a joint cost function $C(y, w, z)$, where y is the vector of n outputs (Moschini 1988). If the regularity conditions of the production process are strengthened by assuming that the input z is normal, I claimed that $\partial^2 C / \partial y_i \partial y_m \geq 0$ for all $i \neq m$. Nelson disputes this result with arguments that revolve around the second-order conditions (SOCs) for problem (15) of Moschini (1988). My own reference to the same SOCs was unnecessary,¹ but the use he makes of them is unacceptable.

Nelson's argument crucially hinges on the claim that the individual cost functions may be concave in

the fixed input, i.e., it is possible to have $\partial^2 C^i / \partial z^i \partial z^i \leq 0$. To understand what this means, it should be clear that $(-\partial C^i / \partial z^i) \equiv r^i(y_i, w, z^i)$ is the shadow price of the fixed input in the i th production line. Hence, Nelson allows this shadow price to be increasing in the quantity of fixed input, i.e., the implicit demand of the fixed input may be upward sloping. One may well wonder under what production conditions this startling feature is admissible. Nelson says that this will happen when the production function is not concave, and he states: "Assuming convexity of all single-output cost functions is the same as assuming that all single-output production functions are concave in z ." Contrary to Nelson's claim, this is not true.

The error in Nelson's argument can be seen by considering the simple counterexample of a Cobb-Douglas production function with only one variable input and one fixed input: $y = xz^2$. This production function is not concave in z , yet the associated cost function is $C = ywz^{-2}$, which is convex in z . In fact, it can be shown that the cost function C^i can be concave in z^i only if the production function violates the property of quasiconcavity. This is seen in the two-input case illustrated in figure 1. If there is only one variable input x and one fixed input z , the shadow price of z can be measured in terms of how much x can be saved as z is varied. In other words, the shadow price of z is the marginal rate of technical substitution multiplied by the variable input price. Hence, for the shadow price to be increasing in z , this marginal rate of technical substitution must be increasing, which means that the technology violates quasiconcavity as is the case for point A in figure 1.

Because this issue is of wider interest, a more rigorous and general proof is warranted. One can prove the following:

PROPOSITION. *If the production function $f^i(x^i, z^i)$ is quasiconcave, the corresponding cost function $C^i(y_i, w, z^i)$ is convex in z^i .*

To prove this, let z_1^i and z_2^i be two distinct levels of the fixed input.² Let x_1^i and x_2^i be the corresponding cost minimizer variable input vector for the output level y_i such that $C^i(y_i, w, z_1^i) \equiv wx_1^i$ and $C^i(y_i, w, z_2^i) \equiv wx_2^i$. Now define $z_0^i \equiv [\lambda z_1^i + (1 - \lambda)z_2^i]$ for $\lambda \in (0, 1)$. Note that, if the production function $f^i(x^i, z^i)$ is quasiconcave, then $[\lambda x_1^i + (1 - \lambda)x_2^i, z_0^i]$ can

Giancarlo Moschini is an associate professor of economics, Iowa State University.

This is Journal Paper No. J-14315 of the Iowa Agriculture and Home Economics Experiment Station, Project No. 2953.

The author thanks Harvey Lapan for his helpful comments.

¹ The SOCs in fact do require the denominator of my equation (18) to be positive. Although justifiable in this sense, reference to the SOCs was unnecessary because, as shown below, standard regularity conditions or a proper definition of input normality are sufficient. In any case, I did not claim, contrary to what is suggested by Nelson, that the SOCs require the denominator of his equation (4) to be positive.

² Note that this proof does not require z^i to be a scalar; hence, the proposition given is valid for the case of many fixed inputs.

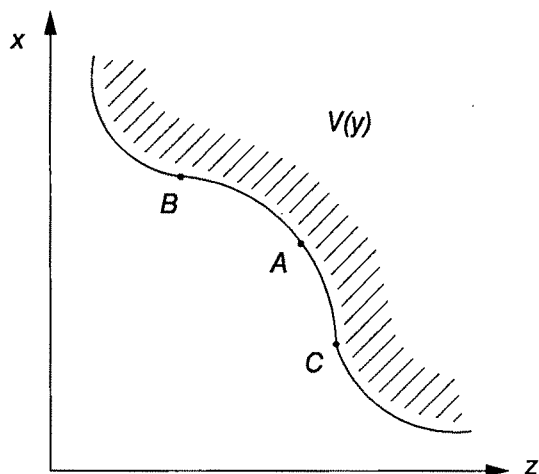


Figure 1. Not quasi-concave technology

feasibly produce y_i (the input requirement set is convex). However, in general this is not the cost minimization solution. Let x_0^i be the cost minimizer when the fixed input is z_0^i . Hence, it must be that $w x_0^i \leq w[\lambda x_1^i + (1 - \lambda)x_2^i]$. By definition $C^i(y_i, w, z_0^i) \equiv w x_0^i$; thus, it follows that:

$$(25) \quad C^i(y_i, w, z_0^i) \leq \lambda C^i(y_i, w, z_1^i) + (1 - \lambda)C^i(y_i, w, z_2^i),$$

which establishes that the cost function is convex in the fixed input. If the production function is strictly quasiconcave, then the cost function $C^i(y_i, w, z^i)$ is strictly convex in z^i . Moreover, if all individual cost functions are convex in z^i , then the joint cost function defined by (15) will be convex in z .³

Nelson's failure to distinguish quasiconcavity from concavity is misleading because quasiconcavity is a much weaker assumption than concavity. Quasiconcavity of the production function implies nonincreasing marginal rates of substitution between inputs and is equivalent to the input requirement set $V(y)$ being convex (the isoquants have the usual convex-to-the-origin shape). It does not imply concavity of the production function, and it does not rule out increasing returns in production (the production possibility set need not be convex). Indeed, the Cobb-Douglas example given above displays increasing returns to scale. When quasiconcavity is violated, some appealing properties of production technologies are ruled out. For example, one may produce a unit of output with the input combinations of either point B or point C in figure 1, but one cannot produce the same unit of output by combining the two production processes

(point on the line segment between B and C are not feasible.)

I have shown that Nelson's argument crucially depends on the violation of the basic regularity condition of quasiconcavity. Because in Moschini (1988) I assumed input normality, I will now argue that for a meaningful definition of a normal input one must be dealing with a quasiconcave technology. The definition of a normal fixed input that Nelson and I use is to require that $\partial^2 C^i / \partial y_i \partial z^i \leq 0$ (for every production line, the shadow price of the fixed input is nondecreasing in output, or, alternatively, the marginal cost is nonincreasing in the fixed input). By itself, however, this is not enough. The standard definition of a normal input requires that the cost-minimizing input demand increases with output (Chambers, p. 69).⁴ Put another way, the output expansion path is not backward bending. When quasiconcavity is violated at the fixed input level z^i , then no expansion path passes through the input allocation point. Thus, the definition of a normal fixed input in terms of $\partial^2 C^i / \partial y_i \partial z^i \leq 0$ is dual to the standard definition only if the technology is quasiconcave. Indeed, the notion of input normality is meant to strengthen the usual regularity conditions, not to weaken them. Hence, I believe that the assumption of quasiconcavity was not only implicit in my analysis but also necessarily implied by my explicit assumption of input normality.

The inconsistency of Nelson's purported assumption of a normal input when quasiconcavity of the production function is violated can be illustrated by a further implication of the case at hand. Recall that input normality in terms of $(-\partial^2 C^i / \partial y_i \partial z^i) = \partial r^i / \partial y_i \geq 0$ implies that the shadow price of the fixed input is nondecreasing in output for all production lines. But what about the shadow price of the same input viewed from the point of view of the whole firm? This shadow price is given by $(-\partial C / \partial z) \equiv r(y, w, z)$, and in equilibrium one must have $r(y, w, z) = r^i(y_i, w, z^{i*})$. This shadow price is simply the negative of the Lagrange multiplier of the problem in my equation (15) [in Nelson's equation (2) this multiplier is θ]. Hence, the question hinges on the sign of $\partial r / \partial y_i = -\partial^2 C / \partial y_i \partial z$. Differentiating the first-order conditions of problem (15) and solving yields

$$(26) \quad \frac{\partial r}{\partial y_i} = - \frac{\frac{\partial^2 C^i}{\partial y_i \partial z^i}}{\frac{\partial^2 C^i}{\partial z^i \partial z^i} \left(\sum_{k=1}^n \frac{1}{\frac{\partial^2 C^k}{\partial z^k \partial z^k}} \right)}.$$

Now, the SOC's for an interior solution of problem (15) allow at most one of the individual cost func-

³ Hence, quasiconcavity of the individual production functions (convexity of the cost functions in the fixed input) means that the curvature conditions for a global minimum for problem (15) are satisfied. One should be careful in using SOC's when the curvature conditions do not necessarily hold because the SOC's only characterize a local optimum. In the present case, if one of the cost functions is allowed to be concave in the fixed input, as Nelson does, there is no guarantee that there exists only one minimum.

⁴ This is the definition given in Moschini (1988). According to another definition, which follows Hicks' discussion of a regressive input, an input is normal if its profit-maximizing demand increases when the output price increases. When the conditions for profit maximization are satisfied, this definition will always imply the previous one.

tions to be concave in the fixed input but require all bordered principal minors of the Hessian to be negative. Note that the n th bordered principal minor in our case is

$$(27) \quad - \left(\prod_{k=1}^n \frac{\partial^2 C^k}{\partial z^k \partial z^k} \right) \left(\sum_{k=1}^n \frac{1}{\frac{\partial^2 C^k}{\partial z^k \partial z^k}} \right).$$

If one of the cost functions is concave in the fixed input, then one must have $[\sum_k (\partial^2 C^k / \partial z^k \partial z^k)^{-1}] < 0$ to satisfy the SOC's. Nelson fails to realize this and gives an incorrect characterization of his own equation (4).⁵ This also means that violation of quasiconcavity for one of the production functions, say the j th, while requiring normality in the sense of $\partial r^j / \partial y_i > 0$ ($\forall i$), implies that $\partial r / \partial y_i < 0$ for all $i \neq j$. Thus, in Nelson's setting one has the paradox of a fixed input that is meant to be normal in all production lines but that is globally inferior in $(n - 1)$ outputs!⁶

Nelson's discussion concerning the relationship of my result to long-run equilibrium conditions is not sound. First of all, because the necessary technological assumption is quasiconcavity, this can relate only to the conditions of long-run cost minimization and not to the conditions of long-run profit maximization discussed by Nelson. More important, it must be recognized that equilibrium is not simply a technological concept and that the properties of the technology provide necessary but not sufficient conditions for long-run equilibrium to hold. For example, if the allocatable fixed input in question is land, a farm may not be at its long-run equilibrium size because of credit constraints, notwithstanding the properties of its production functions (which may, in fact, be quasiconcave). In other words, one cannot infer from observation of disequilibrium that the technological conditions necessary for equilibrium have been violated. In particular, whether or not a fixed input is in equilibrium in the sense of earning the market rate of return is irrelevant for the purpose of understanding the properties of the technology. In the present case, what matters is not whether a fixed input level would be chosen as optimal at the current market rate of return; rather, what matters is whether the level of fixed input could have been chosen as optimal at some appropriate rate of return (i.e., its shadow price).

Nelson's passing comments on Moschini (1989) are similarly without ground. For the sake of clarity, I

will note that in Moschini (1989) I did make the explicit assumption that the individual production functions are concave. This assumption allowed me to generalize the results to profit-maximization conditions and to prove that with allocatable fixed inputs and input normality the supply of any one output is inversely related to the price of any other output.

In conclusion, I have reaffirmed the validity of a result presented in Moschini (1988). Related results were presented in Moschini (1989), and some final thoughts on their relevance are in order. In my analysis I developed the normal case of multioutput production with allocatable fixed factors. An alternative normal case of joint production characterized by Sakai has received considerable attention in the applied agricultural production literature (see, e.g., Hertel, Ball). However, as shown in my work, the case of allocatable fixed factors produces very different conclusions. My results should be of interest for two main reasons. First, my analysis is consistent with the distinctive land allocation problem typical of agricultural production both at the farm and aggregate levels. Second, the assumption that inputs are normal is eminently reasonable for many empirical applications.⁷ Hence, I believe that my results provide a benchmark that is interesting in its own right and that should prove useful in assessing empirical findings and in discriminating among multioutput production structures.

[Received November 1990; no revision.]

References

- Ball, V. E. "Modeling Supply Response in a Multiproduct Framework." *Amer. J. Agr. Econ.* 70(1988):813–25.
- Chambers, R. G. *Applied Production Analysis*. Cambridge: Cambridge University Press, 1988.
- Hertel, T. W. "Estimating Substitution and Expansion Effects: Comment." *Amer. J. Agr. Econ.* 69(1987):188–92.
- Hicks, J. R. *Value and Capital*, 2nd ed. Oxford: Oxford University Press, 1946.
- Moschini, G. "A Model of Production with Supply Management for the Canadian Agricultural Sector." *Amer. J. Agr. Econ.* 70(1988):318–29.
- . "Normal Inputs and Joint Production with Allocatable Fixed Factors." *Amer. J. Agr. Econ.* 71(1989):1021–24.
- Sakai, Y. "Substitution and Expansion Effects in Production Theory: The Case of Joint Production." *J. Econ. Theory* 9(1974):255–74.

⁵ Only one of the three conditions that he cites allow the denominator of his equation (4) to be negative. Moreover, contrary to Nelson's claim, if the j th cost function is the one that is concave in z , it is not sufficient that $\partial^2 C^j / \partial z^j \partial z^j$ be smaller in absolute value than all other such second derivatives to satisfy the SOC's; in fact, its absolute value must be smaller than that when more than two outputs are involved and must decline as the number of outputs increases.

⁶ This is not the only perverse result that one gets when quasiconcavity is abandoned. For instance, although the SOC's allow only one of the individual cost functions to be concave in the fixed input, when this happens it can be shown that the joint cost function $C(y, w, z)$ will be concave in z , despite the fact that $(n - 1)$ individual cost functions are convex in the fixed input!

⁷ Hicks pointed out that it is quite possible that some inputs may be inferior (regressive) because they are better suited for lower levels of production. However, this situation seems unlikely when inputs are aggregated in broad categories such as land, labor, capital, and materials. Because aggregation is typically justified by assuming that input groups are homothetically separable, in this case inputs must be normal within groups as well to guarantee consistent aggregation (expansion paths are straight lines under homotheticity).

Elasticities in AIDS Models: A Clarification and Extension

Richard Green and Julian M. Alston

In a recent paper in this *Journal* (Green and Alston), we showed that the usual formulas for uncompensated price elasticities in the linear approximate almost ideal demand (LA/AIDS) model were incorrect because the Stone's price index (P) is a function of expenditure shares (i.e., $\ln P = \sum_j w_j \ln P_j$). A common approach is to treat expenditure shares as constant parameters in the Stone's index when taking derivatives for elasticities. We developed corrected formulas for uncompensated price elasticities by using derivatives that took into account the effects of price changes on the shares in the price index.

In the process, and as a side-issue, we also noted (p. 443) that the differences in uncompensated elasticities in the literature (η_{ij})

... carry over directly into the computation of compensated elasticities (η_{ij}^*), which are

$$(6) \quad \eta_{ij}^* = \eta_{ij} + w_j (1 + \beta_i/w_i).$$

In making this assertion, we used the usual Slutsky equation in elasticity form and the usual AIDS income (expenditure) elasticity (i.e., $\eta_i = 1 + \beta_i/w_i$). This procedure is correct only when using the true AIDS model.

In the LA/AIDS model, expenditure elasticities also ought to account for the role of expenditure shares as variables in the Stone's price index. That is, the expenditure shares in the Stone's index are functions of both prices and total expenditures unless preferences are homothetic, and expenditure elasticities will be affected along with the uncompensated elasticities (i.e., $\eta_i \neq 1 + \beta_i/w_i$). In general, therefore, correct compensated elasticities cannot be computed simply by substituting correct uncompensated elasticities into equation (6).

In this note we show correct equations for both (a) expenditure elasticities when the LA/AIDS is estimated and (b) compensated price elasticities incorporating those corrected expenditure elasticities.

Richard Green and Julian M. Alston are a professor and an associate professor, respectively, in the Department of Agricultural Economics, University of California, Davis. Senior authorship is not assigned.

Giannini Foundation Paper No. 957.

The authors thank Jim Chalfant and Richard Gray for having pointed out the problem in our previous paper, leading to the present effort. We also thank Giancarlo Moschini and Mike Wohlgemant who offered some additional insights.

Consider the share equation for the LA/AIDS model using our previous notation (Green and Alston):

$$(1) \quad w_i = \alpha_i + \sum_j \gamma_{ij} \ln P_j + \beta_i \ln (X/P).$$

The general expression for the income (expenditure) elasticity in this equation is

$$(2) \quad \eta_{i,x} = d \ln Q_i / d \ln X = 1 + (dw_i / d \ln X) / w_i.$$

The usual approach treats shares as fixed parameters in the Stone's price index (P) and obtains

$$(3) \quad dw_i / d \ln X = \beta_i \text{ and } \eta_{i,x} = 1 + \beta_i / w_i.$$

That is the result used in equation (6) by Green and Alston. More generally,

$$(4) \quad dw_i / d \ln X = \beta_i [1 - d \ln P / d \ln X], \text{ and}$$

$$(5) \quad d \ln P / d \ln X = \sum_j \ln P_j (dw_j / d \ln X) \\ = \sum_j w_j \ln P_j (\eta_{j,x} - 1).$$

Combining equations (4) and (5) gives

$$(6) \quad dw_i / d \ln X = \beta_i [1 - \sum_j w_j \ln P_j (\eta_{j,x} - 1)].$$

Substituting (6) into (2) gives

$$(7) \quad \eta_{i,x} = 1 + (\beta_i / w_i) [1 - \sum_j w_j \ln P_j (\eta_{j,x} - 1)] \text{ or}$$

$$(8) \quad \eta_{i,x} - 1 = (\beta_i / w_i) [1 - \sum_j w_j \ln P_j (\eta_{j,x} - 1)].$$

Following the approach we used for uncompensated price elasticities (Green and Alston), we can solve the system of simultaneous equations in (8) for the elasticities of interest by expressing the equations in matrix form as

$$(9) \quad M = B - BCM,$$

where M is an n -vector with elements $m_i = \eta_{i,x} - 1$, B is an n -vector with elements $b_i = \beta_i / w_i$, and C' is an n -vector with elements $c_j = w_j \ln P_j$. Then, rearranging terms gives

$$(10) \quad M + BCM = (I + BC)M = B.$$

Premultiplying both sides of (10) by $(I + BC)^{-1}$ gives

$$(11) \quad M = (I + BC)^{-1} B.$$

The elements in M are $m_i = \eta_{i,x} - 1$ so we can obtain the expenditure elasticities by adding 1 to each element of M :

$$(12) \quad N = M + \iota = (I + BC)^{-1} B + \iota,$$

where N is an n -vector of expenditure elasticities of demand ($n_i = \eta_{i,x}$), and ι is a unit vector of length n .

Recall the solution for the uncompensated elasticities of demand in the LA/AIDS (Green and Alston, p. 443) is

$$(13) \quad E = [I + BC]^{-1}[A + I] - I,$$

where A is an $n \times n$ matrix with typical elements $a_{ij} = -\delta_{ij} + [\gamma_{ij} - \beta_i w_j]/w_i$ (when $\delta_{ij} = 1$ for $i = j$ is the Kronecker delta) and all of the other variables are as defined above.

Using the Slutsky equation we may find the compensated elasticities of demand as

$$\eta_{ij}^* = \eta_{ij} + w_j \eta_{i,x},$$

and the matrix of compensated elasticities is

$$(14) \quad E^* = E + NW',$$

where E^* is an $n \times n$ matrix with typical elements η_{ij}^* , and W is an n -vector of shares, w_i . Substituting (12) and (13) into (14) yields

$$(15) \quad E^* = [I + BC]^{-1}[A + I] - I + [(I + BC)^{-1}B + \iota]W' \\ = [I + BC]^{-1}[A + I + BW'] - I + \iota W'.$$

Recall that the typical element in A is defined as $a_{ij} = -\delta_{ij} + [\gamma_{ij} - \beta_i w_j]/w_i$. Further, the typical element in I is $+\delta_{ij}$ and the typical element in BW' is given by $(\beta_i/w_i)w_j$. Thus, the term $[A + I + BW']$ reduces to Γ in which the typical element is γ_{ij}/w_i and equation (15) becomes

$$(16) \quad E^* = [I + BC]^{-1}\Gamma - I + \iota W'.$$

To check these manipulations, consider the case when preferences are homothetic (i.e., $\beta_i = 0 \forall i$ or $B = O$, where O is a zero vector). In this case, all of the income elasticities of demand are 1 and this is verified as $N = \iota$ when $B = 0$ in equation (12). In this case ($B = 0$) equation (16) simplifies to

$$(17) \quad E^* = -I + \Gamma + \iota W',$$

and the typical element is

$$\eta_{ij}^* = -\delta_{ij} + (\gamma_{ij}/w_i) + w_j.$$

This last expression is the usual equation for compensated elasticities (as used by Chalfant, for example). The corresponding formula for the elasticities of substitution is

$$(18) \quad \sigma_{ij} = \eta_{ij}^*/w_j = [\gamma_{ij}/w_i w_j] + 1 \quad (i \neq j).$$

Equations (17) and (18) are special cases; in general, they will be incorrect for measuring compensated price responses.

A more general check of equation (16) using homogeneity conditions was suggested to us by Giancarlo Moschini (Iowa State University 1990) in a personal communication. Homogeneity implies $\sum_j \eta_{ij}^* = 0$ or $E^* \iota = O$. We can see that equation (16) satisfies this condition by noting that $\iota W' \iota = I \iota$, and $\Gamma \iota = 0$ when $\sum_j \gamma_{ij} = 0$.¹

In summary, we have extended our results for uncompensated elasticities in the LA/AIDS (Green and Alston) to show corrected formulas for expenditure elasticities and compensated price elasticities that account for the endogenous nature of expenditure shares in the Stone's price index. Equation (7) is the correct equation for LA/AIDS expenditure elasticities. Equation (16) is the correct formula for compensated price elasticities.² Green and Alston's equation (6) is correct for compensated price elasticities of demand only when preferences are homothetic. We regret any confusion we may have caused.

[Received June 1990; final revision received September 1990.]

References

- Chalfant, J. A. "A Globally Flexible, Almost Ideal Demand System." *J. Bus. and Econ. Statist.* 5(1987):233-42.
- Eales, J., and L. Unnevehr. "Beef and Chicken Product Demand." *Amer. J. Agr. Econ.* 70(1988):521-32.
- Green, R., and J. M. Alston. "Elasticities in AIDS Models." *Amer. J. Agr. Econ.* 72(1990):442-45.

¹ Moschini also pointed out another implication of homogeneity. In model (ii) of our earlier paper, the elasticity formula (as used by Eales and Unnevehr, for example) was: $\eta_{ij} = -\delta_{ij} + \gamma_{ij}/w_i$. Homogeneity implies that the Marshallian own-price and cross-price elasticities and the income elasticity sum to zero (i.e., $\sum_j \eta_{ij} = -\eta_{i,x}$), and in the Eales and Unnevehr formula this requires that all income elasticities are unitary ($\eta_{i,x} = 1 \forall i$). Thus, only under homotheticity will the Eales and Unnevehr formula give elasticities that satisfy the homogeneity property, and this accounts for the poor performance of that formula in the empirical results we reported in table 2 since the model we used was not homothetic.

² A referee has pointed out that the corrections to the formulas for calculating compensated elasticities may be unimportant for highly disaggregated commodities for which expenditure shares are so small that compensated elasticities are approximately equal to uncompensated elasticities.

Emerging Issues in the Allocation of Public Agricultural Research Funds

George B. Frisvold

The public and private sectors invest more than \$4 billion in agricultural production research and development (R&D) annually. Adjusting for inflation, the average annual rate of growth of R&D investment has been about 2% since 1970. Federal funding of agricultural research has grown at a considerably slower rate than state and private funding. Thus, the federal government's share of total R&D investment has fallen from 27% to 21% between 1970 and 1987.

At the same time, the traditional emphasis of the public research system on farm-level, production-oriented research has been increasingly challenged on different fronts. The increase in commodity program costs resulting from technical change is receiving greater attention (Alston, Edwards, and Freebairn; Oehmke) as the proportion of net farm income coming from government payments has risen from less than 10% in the 1950s to over 30% in the late 1980s. Representatives from the food industry have called for greater emphasis on post-harvest technologies (Wilcke and Williamson). Over two-thirds of each consumer dollar spent on food goes to the processing and distribution stages of production. Hence, it is argued, greater scope for reducing consumer food costs lies in post-harvest innovations. There is also increased public concern over non-price aspects of food production, particularly issues of environmental pollution, food safety, and nutrition. These concerns imply that demand-side impacts of new technologies are potentially very important.

The public research system faces the following dilemma. On one hand, its constituency has

broadened to include food processors and distributors, input suppliers, and environmental and consumer advocacy groups. On the other hand, fiscal constraints limit the federal government's ability to respond to the demands of its expanded constituency. Consequently, the private sector and state agricultural experiment stations (SAES) will play greater roles in determining the rate and direction of technological change.

These developments raise policy questions regarding the relationship between public research institutions and private industry. For example, to what extent should the public sector cooperate or compete with the private sector in research and product development activities? A related issue is the determination of the optimal mix of basic and applied public research. Another set of policy issues involves consideration of demand-side impacts of new technologies and the coordination of research priorities with other agricultural policy goals. This paper examines these policy issues and addresses their implications for economic analysis of the agricultural research system. A major theme of the paper is the potential for applying theories of industrial organization to address agricultural technology policy questions.

Basic versus Applied Research

Currently, the U.S. Department of Agriculture (USDA) and the SAES allocate approximately 40% of their funds to basic (as opposed to applied) research (Langston). The private sector allocates only about 10% of its research expenditures to basic research (Wilcke and Williamson). Basic research includes attempts to improve basic scientific knowledge without the immediate goal of developing a product for market. Basic research results generally come in

The author is an agricultural economist with the Economic Research Service, U.S. Department of Agriculture.

The opinions expressed are those of the author and are not intended to reflect those of the Economic Research Service, U.S. Department of Agriculture.

the form of new information rather than new commercial products, making it more difficult for researchers to prevent others from imitating or using new discoveries. Basic research results from the public sector can be readily borrowed by the private sector to use in commercial product development. Hence, public basic research acts as a subsidy to the private sector. The applied research and product development by public research institutions directly competes with, rather than complements, private industry. Public applied R&D is also devoted to areas such as disease control for specialty crops or pollution abatement technologies which may be socially desirable but not profitable to develop commercially. An important problem for the public research system is to determine the desirable mix of basic and applied research that avoids duplication of private activities and provides socially beneficial technologies not developed by private industry.

A recent paper by Frisvold models the specific choice of allocating research funds between commodities and between basic and applied research as a simplified portfolio-choice model. An experiment station is assumed to face infinitely elastic input and output markets and must allocate a fixed research budget between two commodities to maximize farm income. For simplicity, two periods are assumed: the present, period 0, and the future, period 1. The agricultural sector's expected future technology in period 1 is characterized by a generalized profit function:

$$(1) \quad \pi = \beta_{11}P_{11} + \beta_{21}P_{21} + f(P_{11}, P_{21}, W_1),$$

where π is discounted profits, P_{11} and P_{21} are the expected prices of the two commodities in period 1, W_1 is a vector of expected input prices, and the function f is nondecreasing in expected output prices and nonincreasing in input prices. The terms β_{11} and β_{21} represent the intercepts of the supply functions for commodities one and two, respectively. Following Binswanger, the β_{i1} terms are functions of the station's research activities such that $\beta_{i1} = \beta_{i1}(R_{i1})$ and $\beta_{i1}' > 0$, where R_{i1} is the future stock of research knowledge. The future research stock is itself a function of current research activities:

$$(2) \quad R_{i1} = R_{i0} (1 - \delta_i) + g_i(R_{i0}, A_{i0}, B_{i0}, I_{i0}); i = 1, 2,$$

where R_{i0} is the current research stock, δ_i is the rate of depreciation, g_i is a concave function, A_{i0} is public applied research expenditures, B_{i0} is

public basic research expenditures, and I_{i0} is the expected level of investment in applied research and product development by private industry. The research station's optimization problem is to maximize π with respect to A_{i0} , A_{20} , B_{i0} , B_{20} given the research budget and its expectations of private investment behavior and future prices. The designation of commodity-specific basic research is justified if the commodity groups are at a high level of aggregation, such as crops and livestock. It is assumed further that basic public research is a complement to both private and public applied research in production of the knowledge stock, while public and private applied research are substitutes. The expected level of private investment in each commodity area is exogenously given. The station must not only allocate research between commodities but also must determine the optimal mix of basic and applied research.

The model yields the following comparative statics results. For commodity i , the optimal ratio of basic to applied research is higher: (a) the greater the level of private investment in commodity i research, (b) the lower the total public research budget, and (c) the lower the initial stock of knowledge in the commodity area. Conversely, a shift to greater emphasis on applied research in commodity area i is warranted: (a) if the private sector does not invest in that area, (b) if commodity i requires high levels of maintenance research (i.e., δ_i is large), and (c) as the price of commodity i increases.

This simplified model makes two very restrictive assumptions. First, the public and private sectors are working to achieve the same types of technological changes at the farm level. Second, no market distortions or environmental externalities are present. Optimal government response to private sector R&D investment, in general, will be highly sensitive to these assumptions. To illustrate, consider a more complete public research objective function, U of the form

$$U = \lambda_f \pi_f + \lambda_i \pi_i + \lambda_p \pi_p + \lambda_c CS - \lambda_g GP - \lambda_e ED - \lambda_r RC,$$

where π_f is farm profits, π_i is input supplier profits, π_p is processor/distributor profits, CS is consumer surplus, GP are government commodity program payments, ED is a dollar value of environmental damage from agricultural production, RC is research costs, and the λ terms are weighting parameters representing distributional preferences of the government toward af-

affected groups. If the government wanted to maximize economic efficiency (given programs), it would equate all λ values. Virtually all private research is carried out by either the input or food industries. Their objectives, rationally, would be to maximize π_i and π_f , respectively. These objectives are not likely to coincide with a government objective, which includes consideration of program and environmental costs of technological change. For example, a significant portion of private R&D is pesticide research for commodity program crops. A government objective function with positive values for λ_g and λ_e would be less likely to complement this research emphasis.

Private industry representatives have argued that the public sector should alter its research emphasis toward more basic research which complements the research activities of post-harvest industries (Wilcke and Williamson). Because the post-harvest phase of production is the area of greatest value added, it is also the area of greatest potential food cost reductions. However, technological change that has the highest payoff to food producers and distributors will not necessarily have the highest benefits to farmers or society in general. Various authors (Alston and Scobie; Freebairn, Davis, and Edwards; Holloway) have demonstrated that it is even possible for post-harvest technological change to reduce farm-level income. These results show that the interests of producers at different stages of the production chain do not always coincide. Thus, the degree to which public institutions will (or should) complement private research agendas depends on how closely those agendas coincide with the preference function, U .

The divergence of interest between the public sector and various agents in the private sector has important implications for assessment of joint public-private research funding. Private grants to state experiment stations represent only about 6% of their total budgets (CSRS). However, a large part of these total budgets are salaries and administrative costs which are relatively independent of the type of research conducted. Private grants, therefore, comprise a more significant portion of a station's discretionary spending. In a case study of Canadian barley research, Ulrich, Furtan, and Schmitz found evidence that private grants from brewers biased the direction of research in their favor and away from the social optimum. Although the social gains from the joint research project were positive, there was

a large public opportunity cost of private influence in the direction of research.

Public and Private Competition

The classic example of direct competition between public research institutions and the private sector has been the seed industry. Public varietal development and releases have been credited with acting as a disciplining device to maintain a relatively competitive structure within the industry. There is growing concern, however, that plant patenting by public research institutions will have a number of detrimental effects. A major argument is that gains from patenting will induce public research systems to behave as profit-maximizing firms rather than social welfare-maximizing ones. Recent literature on the theory of the public firm (De Fraja and Delbono; Cremer, Marchand, and Thisse) calls into question the basic assumption that profit maximization by public firms is suboptimal, however. De Fraja and Delbono use a game-theoretic approach to examine an industry comprised of a public firm and several private, oligopolistic firms. They derive a Nash equilibrium where social welfare is higher when the public firm's goal is profit rather than social welfare maximization. Such a result may or may not apply to experiment stations competing with private seed companies. The point to be made, however, is that new models of optimal public firm behavior may be employed to rigorously evaluate public research strategies.

Product versus Process Research

Cost reduction at the farm level is not always the primary objective of agricultural research. Product quality-enhancing research often affects farm operators more on the demand side of the equation than the supply side. Demand-creating research may be aimed at creating new products or market niches such as those for lower fat or cholesterol foods. Certain new products are not less costly to produce but count on capturing premiums for perceived desirable environmental or health attributes. Examples would be organic produce or "dolphin safe" tuna. Demand-creating research would also include the development of alternative uses for agricultural products or improvement in product desirability to processors or foreign consumers. Other research at-

tempts to alter the maturation date of produce to better take advantage of seasonal price increases. Butler, Carter, and Zepeda's recent study of farmer perceptions of bovine somatotropin also illustrates that farmer technology adoption decisions are increasingly affected by considerations of food safety and consumer acceptance.

Demand-creating R&D may be modeled as an improvement in the average quality of an existing product which increases consumers' willingness to pay for it (Levin and Reiss). Figure 1 demonstrates the welfare effects of an increase in product quality which induces a shift in the demand curve from D to D' . As consumer willingness to pay for the product increases, the equilibrium price rises from P to P' . Producer surplus is increased by the area $PABP'$, while consumer surplus increases by the area $CBP'DAP$. The total gross benefits to product quality improvement is equal to the area $ABCD$.

Government intervention in commodity markets can have important effects on the gains from demand-creating research. Consider the effects of a target price/deficiency payment scheme shown in figure 2. Initially, the target price is set at T , the market price is P , output is Q , government deficiency payments equal the area $TACP$, consumer surplus equals the area PCF , and producer surplus equals the area TAP . An outward shift of the demand curve from D to D' could represent the impact of a technological breakthrough in ethanol production on the demand for corn. In this case, the demand shift has no effect on output, producer surplus, or consumer surplus. The improvement in demand does, however, reduce government payments to

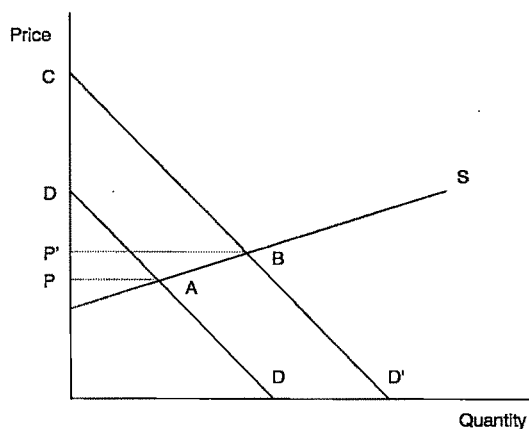


Figure 1. Effects of improved product quality

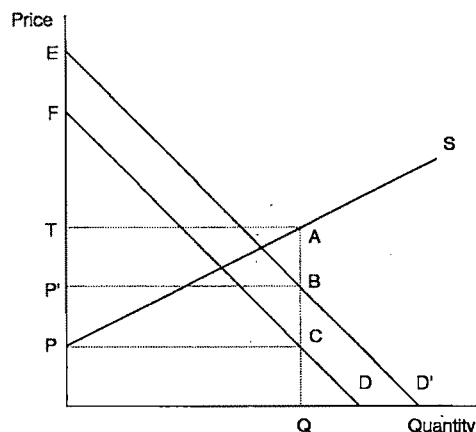


Figure 2. Effect of demand-creating research on a market under a target price/deficiency payment scheme

$TABP'$. Producers have no incentive to fund such demand-creating research because it would only change the proportion of their income coming from the market as opposed to that coming from the government, leaving total returns unchanged.

Three important points can be made here. First, demand-creating innovations can reduce the social costs of farm income support policies. In contrast, cost-reducing innovations increase the cost of target price programs. Second, target price schemes can make demand-creating innovations less attractive to producers than cost-reducing innovations. Finally, to the extent that consumers are also taxpayers, they may benefit more from demand-creating innovations, which reduce the tax burden of income support programs, than from cost-reducing innovations, which lower food costs but raise their tax burden.

Product quality research is complicated by problems of imperfect information. Many quality characteristics consumers care about, such as flavor, are not readily observable at the time of purchase. Consequently, consumers use appearance as a signal of underlying product quality. The signal value to consumers of product appearance (and hence the premium they are willing to pay) depends on the degree of correlation between appearance and unobserved quality.

A recent paper by Frisvold and Putler examines the welfare effects of research on product attributes in the context of asymmetric information concerning the relationship between

unobserved product quality and appearance. Consumers select a product given its price and expected quality given its appearance, updating their expectations of the quality-appearance correlation given new information. The producer side represents an experiment station or marketing order that invests in research to alter product appearance and/or unobserved quality to maximize net producer returns to research. Researchers face the following intertemporal tradeoff. In the short run, there are gains from altering product appearance only, while the gains from improving average quality accrue only in the long run. Reducing the appearance-quality correlation, however, erodes the signal value of appearance over time. Examples of such "signal jamming" effects have been observed in the markets for tomatoes and stone fruits.

The imperfect information equilibrium is characterized by overinvestment in appearance and underinvestment in quality relative to the social optimum. A net social loss from research is possible such that surplus is simply transferred from consumers to producers while research expenditures represent a type of dead-weight loss. Key variables driving model results are consumers' marginal utility with respect to quality, researcher and consumer expectation formation, and the researcher's discount rate. The model has two major implications for grading standards. First, grading standards based on appearance alone can have negative long-run welfare effects. Second, there are potential gains from improved grading/labelling based on quality which is unobservable to consumers. Such standards not only reduce consumer uncertainty but can create research incentives more compatible with the socially optimal path of technological change.

Conclusion: Future Areas of Research

Historically, major technological breakthroughs have come from publicly funded research on farm-level innovations. In the future, such breakthroughs are more likely to come from privately funded research on post-harvest innovations. As a consequence, traditional methods of analyzing the impact of public research in terms of simple supply shifts while ignoring private sector behavior will prove increasingly inadequate. Industrial organization theories of imperfect and non-price competition, the public firm,

and asymmetric information can play an important role in addressing emerging policy issues in agricultural research.

The refinement and empirical application of these theories to agriculture implies a need for much better information on private R&D activities. Such data are vital in assessing the impact of government policies toward basic research, scientific education, intellectual property rights, antitrust policy, and investment tax credits on the extent and direction of private innovation. Greater knowledge of private R&D trends is essential if the public sector wishes to reduce redundancy in research and to allocate funds to higher payoff investments. Ironically, the availability and quality of private R&D data appears to be declining. Unless institutional innovations are forthcoming that reverse this growing information deficit, policy makers will be less able to identify those factors that retard or promote technological change.

Finally, there is a need for more empirical analysis of the behavior of public research institutions. Two possible areas of inquiry include patenting by public institutions and joint public-private research ventures. A positive theory of government behavior is necessary for a complete understanding of the process and direction of technical change. Revealed preference estimation of government technology policy may also be a fruitful avenue of research.

References

- Alston, J. M., and G. M. Scobie. "Distribution of Research Gains in Multistage Production Systems: Comment." *Amer. J. Agr. Econ.* 65(1983):353-56.
- Alston, J. M., G. W. Edwards, and J. W. Freebairn. "Market Distortions and Benefits from Research." *Amer. J. Agr. Econ.* 70(1988):281-88.
- Binswanger, H. P. "Issues in Modeling Induced Technical Changes." *Induced Innovation: Technology, Institutions, and Development*, ed. H. P. Binswanger and V. W. Ruttan. Baltimore MD: Johns Hopkins University Press, 1978.
- Butler, L. J., H. O. Carter, and L. Zepeda. "Adoption and Diffusion of New Technologies: Bovine Somatotropin and the California Dairy Industry." Dep. Agr. Econ. interim rep., University of California, Davis, 1990.
- Cooperative State Research Service (CSRS). *Inventory Agricultural Research*. Washington DC, various years.
- Cremer, H., M. Marchand, and J. Thisse. "The Public Firm as an Instrument for Regulating an Oligopolistic Market." *Oxford Econ. Pap.* 41(1989):283-301.

- De Fraja, G., and F. Delbono. "Alternative Strategies of a Public Enterprise in Oligopoly." *Oxford Econ. Pap.* 41(1989):302-11.
- Freebairn, J. W., J. S. Davis, and G. W. Edwards. "Distribution of Research Gains in Multistage Production Systems: Reply." *Amer. J. Agr. Econ.* 65(1983):357-59.
- Frisvold, G. B. "Public Agricultural Research Investment Decisions: Basic vs. Applied." Mimeographed. Washington DC: U.S. Department of Agriculture, Economic Research Service, 1990.
- Frisvold, G. B., and D. Putler. "Product Quality Research Under Imperfect Information." Mimeographed. Washington DC: U.S. Department of Agriculture, Economic Research Service, 1990.
- Holloway, G. J. "Distribution of Research Gains in Multistage Production Systems: Further Results." *Amer. J. Agr. Econ.* 71(1988):338-43.
- Langston, J. B. "Dynamic Strategies for Research Expenditures in Agriculture: Public and Private, Basic and Applied." Ph.D. thesis, University of Georgia, 1988.
- Levin, R. C., and P. C. Reiss. "Cost-Reducing and Demand-Creating R&D with Spillovers." *Rand J. Econ.* 19(1988):538-56.
- Oehmke, J. F. "The Calculation of Returns to Research in Distorted Markets." *Agr. Econ.* 2(1988):289-302.
- Ulrich, A., H. Furtan, and A. Schmitz. "Public and Private Returns from Joint Venture Research: An Example from Agriculture." *Quart. J. Econ.* 101(1986):104-29.
- Wilke, H. L., and J. L. Williamson. *A Survey of U.S. Agricultural Research by Private Industry*. Washington DC: Agricultural Research Institute, 1977.

Plant Variety Protection, Private Funding, and Public Sector Research Priorities

Mary K. Knudson and Carl E. Pray

The objective of the Plant Variety Protection Act of 1970 (PVPA) was to increase a firm's ability to capture the returns on its applied breeding research and development (R&D) investments. Firms were expected to increase their R&D investments to take advantage of increased returns. Studies have shown that the PVPA has induced the private sector to increase its investments, particularly in soybeans and wheat (Butler and Marion; Perrin, Kunnings, and Ihnen). Recently, Stallman raised the concern that the PVPA may also influence agricultural experiment stations (SAES). She suggests that public R&D priorities may be shifted away from commodities and research topics that have high social rates of return.

While it is possible that the PVPA will pull research away from the socially optimal allocation of resources, other sources of distortion may be equally important. Political pressure on public research or industry's contributions to public research may lead to too much research on certain topics (Ulrich, Furtan, and Schmitz) or commodities. Another concern is that, in certain commodities, the private sector research is crowding out public research. Both concerns raise the questions of who will teach the next generation of scientists, who will provide varieties when markets are insufficient to induce private research, and who will maintain the competitive structure of agricultural production, inputs, and marketing.

This paper is part of a larger project on the impact of the PVPA and patents on public and private R&D, and on crop yields and genetic

diversity of major field crops in the United States. The primary purpose of this paper is to measure the PVPA's impact on public sector plant breeding. The secondary purpose is to find out whether industry directly influences public research. As a preliminary study, we examine public sector R&D expenditure patterns on wheat, soybeans, corn, cotton, and sorghum.

Public Sector Resource Allocation and Intellectual Property Rights

Table 1 shows why some people are concerned. Since 1980 the number of scientists years (SYs) in some of the Agricultural Research Service (ARS) and SAES plant breeding programs have declined. In corn, SYs dropped 22%, the most of the five crops. However, crowding out is most likely to occur in corn because corn attracts more private resources and is growing at a faster rate than the other four crops. Sorghum, too, has a long history of private research and experienced a decline in public sector SYs. In contrast public SYs have increased in soybeans and wheat, which is where the PVPA strengthened property rights the most. Cotton is an anomaly; it declined even though the PVPA strengthened SAES' ability to appropriate gains.

Table 2 shows that SAES have been acquiring Plant Variety Protection Certificates (PVPCs). SAES started applying for wheat, soybean, and cotton PVPCs in 1970. The first two columns in table 2 show that the public sector acquired the greatest number of PVPCs in soybeans and wheat. No PVPCs were issued for corn or sorghum inbred lines even though these crops were eligible for protection.

The second half of table 2 shows the results of a 1989 ARS survey of SAES on their use and future intentions to use PVPCs and utility patents (UPs). Two-thirds of the SAES currently use PVPCs, but only two more plan to use them

Mary K. Knudson is a visiting assistant professor in the Institute of Public Policy Studies at the University of Michigan. Carl E. Pray is an associate professor in the Department of Agricultural Economics, Rutgers University. Authorship is shared equally.

This research was partially supported by a cooperative agreement between Rutgers University and the Economic Research Service, U.S. Department of Agriculture. The views and errors in this paper are the authors'.

The authors would like to thank Glenn Nelson, Jim Oehmke, John Reilly, and Marie Walsh for their comments.

Table 1. Public and Private Scientists in Applied Research on Selected Crops

	Public				Private	
	68	72	80	88	82	89
Corn	56	58	68	53	155	257
Sorghum	12	16	17	14	22	23
Wheat	39	54	58	69	23	25
Cotton	50	43	41	37	17	11
Soybeans	18	26	42	59	36	60

Source: Public SYs from CRIS; Private Ph.D. SYs from Kalton, Richardson, and Frey.

Table 2. PVPA Certificates and Patents Issued and Intentions of Public Sector for Selected Crops, 1973-89

Crop	PVPCs Issued to SAES or ARS		No. of SAES which:			
	1973-81	1982-89	Use PVPA ^a	Will Use ^b	Use Patents ^a	Will Use ^b
Corn and sorghum	0	0	1	5	3	5
Cotton	12	7	4	6	1	5
Soybean	27	59	14	19	1	5
Wheat	29	26	11	17	0	4
All crops			32	34	19	37

Sources: Columns 1 and 2, "Plant Variety Protection Office Journal." Columns 3-6, Howard J. Brooks, "Questionnaire on Maintenance of Free Exchange of Plant Germplasm," USDA ARS Memorandum, 17 Nov. 1989.

^a The number of AES that applied for PVPCs or patents as of 1989.

^b The number of AES that stated that future use of plant protection was either definite or probable.

in the future. The number of SAES that plan to use UPs will almost double from nineteen to thirty-seven. The number of SAES using PVPCs and UPs for the major field crops, particularly corn, sorghum, and wheat, will grow more rapidly than the totals. At present only a few experiment stations use UPs on these crops.

Model of Public Research Resource Allocation

Most economists advise government research administrators to allocate their research resources to maximize expected social benefits from research. But most empirical studies of research resource allocation in the agricultural economics literature assume that research administrators are also influenced by the demands of organized interest groups (Huffman and Miranowski, Hayami and Ruttan, Rose-Ackerman and Evenson). These models assume that research administrators are cost minimizers within a budget that is partly determined by interest groups and politicians's demand for research.

An alternative way to model the supply side is to consider ARS and SAES administrators as

budget-maximizing bureaucrats (Ruttan). The director invests his research resources in different commodities to increase his research resources in the future. Major sources of funding are state and federal governments, private industry and royalties from PVPCs and UPs. Under this scenario, for each funding source, administrators will like to increase its research contributions. Because the funding sources vary with regard to their own research objectives, the strategies that the administrator employs to gain this increase will vary depending on the funding source.

To increase state and federal contributions, the ARS and SAESs could either develop technologies that reduce farm production costs or that increase consumer welfare and the income of influential farmers and agribusiness. Thus, the director would want to maximize total social benefits. Measurements of recent research productivity and possible social benefits from this research could help determine where these benefits would be highest. In addition, the director might reduce expenditures in commodities where private sector research is strong because social benefits from duplicating private research would be limited.

Research directors would probably take a different approach to increase industry's contributions. For example, the ARS and SAESs could concentrate on research that increases the profits of agribusiness and commodity groups. More immediate increases might come if the public sector shifted resources onto topics which the private sector already helps fund.

Since the passage of PVPA, public research institutions, along with firms, can collect royalties from sales of pure line varieties of sexually propagated crops. To take further advantage of income from royalties, the ARS and SAESs may seek more market space by developing superior varieties to those that the private sector produces.

In our model, ARS and SAES administrators will consider social benefits from research, contributions from industry, and royalties from PVPCs and UPs when allocating research resources among commodities. More formally, this relationship can be written as

$$(1) \quad RD_{ij} = f(SB_{ij}, NPR_{ij}, IND_{ij}, PRIV_{ij}, RP_{ij}),$$

where RD is ARS and SAES expenditure at time i on crop j , SB is the expected social benefit from research, NPR is expected royalty income to SAES from research, IND is industry's funding of public research, $PRIV$ is the quantity of private research, and RP is a measure of public research productivity.

The actual variables used and sources of data are shown in table 3. However, some of the independent variables need further explanation. One common way of measuring social benefits is net social surplus (NSS) (Perrin and Foster):

$$(2) \quad NSS_j = c * k_j * V_j * [1 + k/(2/n_j + 2/e_j)] - C_j(k),$$

where for every crop j , c is a capitalization factor (which captures the present value of social benefits), V is the value of production, k is the percent reduction in production costs resulting from a new technology, e is the elasticity of supply, n is the absolute value of the elasticity of demand, and C is R&D costs. We assume that c is the same for all crops, so $c = 1$. For simplicity, we also assume that the reduction in production costs is 1% for all five crops, so $k = 1$. Therefore, equation (2) becomes

$$(3) \quad SB_j = V_j * [1 + 1/(2/n_j + 2/e_j)] - C_j,$$

which we use to estimate SB_{ij} . Several different lag structures on the value of production were used in estimating SB_{ij} . However, an average over the previous five years proved to be the best lag.

Experiment station directors might also look at measures of the productivity of research programs other than k such as numbers of varieties released or publications. We tried as a rough measure of research productivity (RP) the number of varieties released divided by earlier research expenditure. The coefficients for all variables, with the exception of $PRIV$, are expected to be positive.

Empirical Results

Three specifications of our model [equation (1)] were estimated in linear form using OLS. Several functional forms of our model were estimated, but the linear form had the best fit. The results of the linear specifications are shown in table 4. Social benefits (SB) are consistently positive and significant at the 5% or 10% level across all three specifications. Research produc-

Table 3. Variables and Data Source

Variables	Description and Data Source
Dependent variables	
<i>RD</i>	Federal and SAES R&D plant breeding and maintenance (RPAs 307 and 405) from CRIS data in time periods 1968, 1972, 1976, 1980, 1984, and 1988 and crops corn, wheat, soybeans, cotton, and sorghum.
Expected social benefits	
<i>SB</i>	Producer and consumer benefit calculated using Perrin and Foster (see text).
<i>RP</i>	Productivity of research for i less than 1980 is the number of varieties for each j released in <i>Crop Science</i> 1968–70 divided by the number of SYs in 1968. For i greater than or equal to 1980, the number of varieties for each j from 1985–87 was divided by the number of scientific years in 1985.
<i>PRIV</i>	Private expenditure on research from Perrin, Kunnings, and Ihnen; and Kalton, Richardson, and Frey. This variable was lagged in the following manner: for 1968, 1972, 1976, 1980, 1984, and 1988; private expenditures were taken from 1960, 1965, 1970, 1975, 1980, and 1982, respectively.
SAES income generation	
<i>NPR</i>	Dummy for the PVPA is one for the self-pollinated crops in which SAES took PVPCs—wheat, soybeans, and cotton—after 1970.
<i>IND</i>	Industrial funding per crop per year from CRIS data.

Table 4. Factors Influencing Public Allocation of R&D Resources: Regression Results

Dependent Variable: Applied Public R&D by Commodity			
Independent Variables	Specifications		
	1	2	3
Social benefits (<i>SB</i>)	0.00011 (1.64)	0.00034 (6.884)	0.00013 (1.91)
Research productivity (<i>PROD</i>)	-0.00003 (-.059)	0.0004 (.740)	0.00011 (0.25)
Private research (<i>PRIV</i>)	-0.030 (-.93)		
Potential royalties (<i>NPR</i>)		0.0017 (2.724)	0.00075 (1.38)
Industry contributions (<i>IND</i>)	6.40 (4.93)		5.22 (3.91)
Constant	0.0009 (2.04)	0.0003 (.500)	0.0006 (1.07)
Adjusted R^2	.828	.648	.77
Degrees of freedom	25	26	25

tivity (*RP*) changes signs and is not significantly different from zero in any specification. Private research (*PRIV*) has the expected sign but is not significant. However, this relationship deserves a more in-depth examination which is beyond the scope of this paper. The intellectual property rights coefficient (*NPR*) is positive in most specifications but is statistically significant only in specification 2. When the industry contributions variable (*IND*) is added (specification 3) the R^2 increased, the *NPR* coefficient remained positive but was no longer significant, and the *RP* coefficient changed to the expected sign. *IND* is positive and significant in all specifications of the model.

The regression results support the argument that expected social benefits guide research directors in their allocation of resources. They also suggest that intellectual property rights have influenced the commodity composition of public research. They do not support the argument that research contributions by industry and commodity groups induce directors to allocate more government resources to those commodities. The industry variable is positive, but the elasticity of public research with respect to industry contributions to public research is less than one (.5 to .6, depending on the specification). This result indicates that public research administrators substitute private contributions for public research dollars rather than spending more public research dollars to capture limited industry contributions, as some critics predicted.

Conclusions

The regression results indicate that social benefits are important in influencing the direction

of research. There is some support for the argument that the new opportunities for income provided by the PVPA have influenced the direction of public research. The key policy question, however, is whether PVPA and industry contributions would result in lower social benefits.

The first column in table 5 shows the most recent (1988) allocation of public resources between the five crops in this study. If in the future resources are allocated to crops in which property rights have been improved, wheat, cotton, and soybeans would receive more resources. If public resources were allocated by the value of the seed market, they would be allocated like column 2. The share of corn would more than double, wheat's share would be cut in half, and cotton's share would decline from 14% to 3%. If resources are allocated more like contributions of private companies and commodity groups (column 3), the corn and soybeans will increase while the other crops' shares will decline. If the public sector tries to duplicate the private sector (column 4), even more emphasis would be placed on corn.

Would reallocating the public sector resources according to the value of the seed market reduce social benefits? A definitive answer is not possible here. However, if resources were allocated according to our crude measure of social benefits, they would resemble column 5. This column suggests that public resources would shift away from wheat and cotton toward corn. Thus, if experiment stations shifted resources in the direction suggested by industry's contribution, they would also be moving toward an allocation of resources that would increase social benefits. If they allocate resources according to the value

Table 5. Current and Potential Future Allocation of Research Resources—Selected Crops (percent of 5-crop total)

	Current (1988) Public Allocation	Allocation Based on Seed Market	Industry Contribution	Private R&D 1989	Allocation Based on SB
Corn	25	58	34	68	40
Sorghum	8	6	6	6	4
Wheat	30	12	19	7	18
Cotton	14	3	11	3	9
Soybeans	24	21	29	16	29

Sources: Column 1, CRIS; column 2 calculated from value of commercial seed in mid-1980s from Daberkow; column 3 calculated from CRIS data for 1988; column 4 calculated from Kalton, Richardson, and Frey; column 5 calculated from social benefits for 1988; see table 4.

of the seed market or like the private sector, they might be placing too much emphasis on corn.

References

- Butler, L. J., and B. W. Marion. *The Impacts of Patent Protection on the U.S. Seed Industry and Public Plant Breeding*. North-Central Regional Research Publication 304, Sep. 1985.
- Daberkow, S. *Agricultural Input Industry Indicators in 1975–85: Expansion and Contraction*. Washington DC: U.S. Department of Agriculture, Economic Research Service Agr. Info. Bull. No. 534, Nov. 1987.
- Hayami, Y., and V. W. Ruttan. *Agricultural Development: An International Perspective*. Baltimore MD: Johns Hopkins University Press, 1985.
- Huffman, W. E., and J. A. Miranowski. "An Economic Analysis of Expenditures on Agricultural Experiment Station Research." *Amer. J. Agr. Econ.* 63(1981):104–18.
- Kalton, R. R., P. A. Richardson, and N. M. Frey. "Inputs in Private Sector Plant Breeding and Biotechnology Research Programs in the USA." *Diversity* 5(1990):22–25.
- Perrin, R. K., and W. Foster. "Economic Measures of the Effect of Intellectual Property Rights on Plant Breeding." Paper presented at the AAAS meetings, New Orleans LA, 18 Feb. 1990.
- Perrin, R. K., K. A. Kunnings, and L. A. Ihnen. "Some Effects of the U.S. Plant Variety Protection Act of 1970." Dep. Econ. and Bus. Econ. Res. Rep. No. 46, North Carolina State University, Aug. 1983.
- Rose-Ackerman, S., and R. E. Evenson. "The Political Economy of Agricultural Research and Extension: Grants, Votes, and Reapportionment." *Amer. J. Agr. Econ.* 67(1985):1–14.
- Ruttan, V. W. "Bureaucratic Productivity: The Case of Agricultural Research." *Public Choice* 35(1980):529–47.
- Stallman, J. I. "Plant Patents and Public Research Priorities." *Choices* (1990):8–11.
- Ulrich, A., H. Furtan, and A. Schmitz. "Public Returns from Joint Venture Research: An Example from Agriculture." *Quart. J. Econ.* 100(1986):103–29.

Inventions Intended for Use in Agriculture and Related Industries: International Comparisons

Robert E. Evenson

Private firms invest substantial funds in research and development (R&D) to discover and develop inventions intended for use by crop and livestock farmers.¹ Most of these inventions are granted patent protection, and many inventors obtain patent protection in more than one country for their inventions. Publicly supported agricultural R&D, including research done in the U.S. Department of Agriculture-State Agricultural Experiment Station (USDA-SAES) system, also produces inventions intended for use by farmers, but few of these inventions are granted patent protection.²

A recently completed research project conducted at Yale University provides procedures for assigning patented inventions to "Industries of Use" (IOUs) including agricultural production industries.³ Specifically, the project developed a "concordance" mapping inventions classified by International Patent Class (IPC) into industries classified by Standard Industrial Class (SIC). The concordance can be applied to international patent data sets reporting IPCs and thus provides estimates of inventions by IOU (or by Industries of Manufacture, IOM) by country of origin.⁴ This paper reports estimates of inventions intended for use in U.S. agricultural production sectors as well as in the input supply and agricultural service sectors. Data for Great

Britain, France, and Germany are utilized for international comparisons and for estimates of technology flows between countries.

Part 1 discusses the concordance methodology. The second part reports estimates of inventions by two- and three-digit industries of use. The third section reports international comparisons.

The Yale-Canada Concordance

The concordance utilizes data from the Canadian Patent Office to compute probability distributions mapping inventions from an International Patent Class to two-, three-, or four-digit SIC Industries of Manufacture and Industries of Use.⁵

The Canadian Patent Office has been assigning an IPC, an IOM, and an IOU to each patent granted since 1978 (to Canadian origin patents since 1972). (See Ellis for a full discussion.) This provides a large data base (more than 200,000 patents) for an international sample of inventions (only 12% of Canadian patents are of Canadian origin, nearly half are granted to U.S. inventors). The IPC structure is very detailed, with roughly 100,000 classes. (The actual concordance is based on a more aggregate set of 6,000 classes.) The procedure to estimate the concordance is to compute the distribution across IOMs or IOUs for each detailed IPC for the Canadian sample. The reliability of this concordance, when applied to other patent data sets (e.g., for the U.S.) depends on the inherent "diffuseness" or matching of IPCs with IOUs or IOMs and/or possible changes in industry mixes. Tests of reliability (Kortum and Putnam) of the concordance showed that the concordance was

Robert E. Evenson is a professor, Economic Growth Center, Yale University.

¹ Evenson and Huffman, report estimates of R&D expenditure by private firms in the U.S. farm implements industry of \$592 million in 1984 and of \$837 million by firms in the agricultural chemicals industry.

² See Evenson and Putnam, and Evenson and Huffman for a fuller discussion of protection of USDA-SAES inventions.

³ R&D data are reported by the industrial firms conducting the R&D. One can infer only roughly the industry of use for this R&D and for the associated inventions.

⁴ Evenson and Huffman, chap. 7, and Evenson and Putnam also report data on inventions by technology field for agriculture. They note a growing proportion of inventions granted to foreign investors in all agriculturally related technology fields in recent years.

⁵ See Kortum and Putnam for an extended analysis of the concordance and an alternative concordance.

very reliable at the two-digit level and generally reliable at the three-digit level.⁶

Given the uniqueness of these data, however, it will be useful to generate data for some four-digit sectors and to interpret them in the context of their limited reliability. (Note the estimated series are unbiased estimates of inventions by IOU.) In the sections to follow, estimated series by IOU have been generated from the U.S., U.K., French, and German patent data bases of the European Patent Office (EPO). For the United States, the EPO has assigned IPCs to U.S. patents for the 1920–69 period, and these historical data are summarized in this paper as well.⁷

Inventions by Industry of Use

The Farm Industries

Farms conduct very little R&D, and farmers make few inventions (Evenson and Huffman, chap. 5). Farm industries, however, are the industries

of use for many inventions. During the 1970s more than 900 inventions per year granted patents by the U.S. Patent Office were destined for use in the farm sector. More than 25% of these were “imported” from abroad, i.e., they were invented in other countries.⁸

Table 1 reports an historical summary of inventions intended for use in six three-digit SIC farm industries with details for selected four-digit SIC industries. It should be noted that many inventions assigned to the three-digit class are not further assigned to a four-digit class. Thus, for each four-digit industry, one may view these inventions as specific to the industry in question; but further note that the industry also uses inventions that are specific to the three-digit industry.

The first seven columns in table 1 report historical data for all inventions granted patents in the United States, i.e., including inventions originating in foreign countries (see table 3 for further details on countries of origin). Columns eight and nine report import ratios, i.e., the ratio of imported inventions to domestic inventions, for the 1970s and 1980s. These import

⁶ Tests are made (Kortum and Putnam) using part of the Canadian data to predict distributions for the remaining data. No direct tests of the international applicability of the concordance were made.

⁷ The International Patent Office assigned IPCs to U.S. patents back to 1920, enabling the historical series. The U.S. Patent Office did not make such assignments until 1965.

⁸ Evenson and Huffman report estimates of inventions used in agriculture based on “technology fields.” Their estimates show only one-third as many inventions as reported here.

Table 1. Inventions Used by the U.S. Agricultural Production Sector

	Average Annual Inventions Patented							Import Ratio		Export Ratio 1970s
	1920s	1930s	1940s	1950s	1960s	1970s	1980s	1970s	1980s	
011 Livestock farms (total)	204.4	166.3	96.6	178.3	211.3	230.9	202.3	.21	.40	.42
0111 Dairy farms	25.0	22.5	14.3	21.7	23.8	28.5	23.8	.28	.55	.44
0112 Cattle farms	13.0	10.4	7.6	17.2	23.3	27.8	25.9	.20	.41	.39
0113 Hog farms	6.1	8.5	2.6	8.2	10.9	9.7	10.0	.09	.48	.49
0114 Poultry farms	76.9	54.8	25.4	50.0	44.6	38.9	34.3	.17	.31	.34
0115 Sheep-goat farms	.1	.1	.2	.3	.5	.7	1.0	.90	.13	.03
012 Other animal farms (total)	39.7	23.3	14.2	20.5	19.7	22.6	26.7	.18	.19	.17
0121 Beekeeping farms	13.4	8.9	5.5	7.2	5.4	6.4	10.5	.09	.22	.14
0122 Horse farms	18.2	7.6	3.7	5.6	6.8	8.2	8.4	.26	.14	.11
0123 Fur skins farms	.7	.2	.1	.4	.3	.3	.2	.00	.00	.00
013 Field crop farms (total)	443.6	320.4	227.3	425.7	383.8	406.7	332.0	.26	.47	.58
0131 Wheat farms	32.5	23.8	13.9	23.4	29.0	33.3	29.2	.27	.46	.40
0132 Small grains farms	4.4	5.9	7.0	11.1	15.9	15.3	10.6	.47	.47	.55
0134 Corn farms	8.9	7.3	7.5	11.5	9.9	10.6	9.1	.16	.52	.37
0135 Forage-hay farms	8.8	6.2	6.1	18.1	17.4	15.4	16.7	.26	.58	.59
0137 Tobacco farms	1.0	1.4	1.1	2.0	3.5	3.5	4.6	.13	.78	.58
0138 Potato farms	6.8	4.4	3.3	6.0	3.4	3.6	2.6	.20	.21	.43
015 Fruit and vegetable farms	43.8	37.9	23.1	57.5	68.2	64.9	35.7	.15	.31	.21
016 Horticultural farms (total)	23.0	31.9	16.4	32.6	37.7	56.3	57.1	.21	.33	.28
0161 Mushrooms	1.7	2.8	1.5	4.2	4.6	4.8	4.7	.17	.40	.27
0162 Greenhouse products	6.2	9.2	4.3	6.7	8.1	10.7	11.0	.26	.30	.34
0163 Nursery products	2.2	2.7	1.4	3.4	4.4	7.4	6.2	.21	.27	.25
017 Combination farms	68.1	67.3	57.5	106.5	138.4	142.4	150.2	.35	.60	.60
All farms	824.6	647.1	435.1	819.1	859.1	923.8	804.0	.25	.45	.50

ratios changed dramatically from the 1970s to the 1980s.⁹

The historical series show quite regular patterns. Inventions for all farm sectors (except for horticultural farms) declined during the 1930s and 1940s, then rose sharply in the 1950s, 1960s, and 1970s. Most sectors show a decline in total inventions in the 1980s, and all sectors show a decline in domestic origin inventions in the 1980s. The decade of the 1930s was affected by the Great Depression, and the 1940s were affected by World War II. Most of the inventions destined for use in the farm sectors originate in the industries supplying agriculture (see table 2), and these industries grew in importance in the 1950s and 1960s.

Comparisons across four-digit industries have limited relevance because many inventions can be used in several four-digit industries and hence are assigned only at the three-digit level. Such comparisons are of some interest, however. For example, they show differences in the timing of inventions, e.g., the early boom in poultry inventions in the 1950s and high levels of inventions in the 1980s in specialized animal and nursery farm industries. Comparisons across three-digit industries show that field crop farms have received more inventions than livestock farms.

Returning to the matter of the import ratios, it is clear that these ratios differ between industries, being lowest for livestock, fruit and vegetable, and horticulture farms and highest for field crop farms. These ratios reflect two factors. First, there are fundamental differences in technologies and their transferability. Some inventions are highly location specific and hence worth little in another country. Second, increases in the import ratios from the 1970s to the 1980s indicate that the "competitiveness" of foreign inventors has risen markedly in the 1980s. Not only has the foreign share risen, but domestic invention has declined sharply for most industries (specifically for the field crop farm industries).¹⁰

Table 1 also reports an export ratio, i.e., the ratio of patents granted to U.S. inventors in France, Germany, and the United Kingdom to U.S. domestic patents. This ratio has remained

approximately constant over the 1970s and 1980s. While not fully comparable to the import ratio (which includes patents from all foreign countries), it does show that U.S. inventions have been competitive abroad (see section 3 below for further discussion).¹¹

The Farm Supply Industries

Table 2 reports data comparable to table 1 where the industries of use are the farm supply industries. These industries are the primary sources of inventions for the farm industries. Table 2 is designed to show the flows of inventions to these supplying industries and perhaps to shed some light on the apparent loss of competitiveness of these industries as invention suppliers to the farm industries.

The same general historical patterns exhibited in table 1 appear for the industries in table 2. Inventions used in these industries declined during the 1930s and 1940s and rose over the 1950s, 1960s, and 1970s. Total inventions declined for both the agricultural implements and the agricultural chemicals sector in the 1980s, and the rising import ratios indicate a sharp decline in domestic invention. The 27% decline in inventions of use for the agricultural implements and agricultural chemicals industries from the 1970s to the 1980s was considerably greater than the 15% decline for the farm industries for this period.

The services sector industries are the industries of use of significant numbers of inventions. These industries actually perform many farm functions. As with other sectors these too have experienced increased foreign competitiveness.

International Comparisons

Table 3 reports comparative import and export ratio data for the four major origin countries for agricultural inventions. It also reports invention by source country for the 1969 to 1987 period. These comparisons are reported for the major three-digit industrial sectors of use including crop and livestock farms.

For each industry two further indexes are re-

⁹ Evenson and Huffman report data by technology field, indicating that the import ratio for all agriculturally related inventions rose from less than .2 in the 1920s and 1930s to .4 by the 1950s and approximately .5 by the 1960s and 1970s.

¹⁰ The question of rising competitiveness has substantial policy relevance and requires further extensive analysis. (See Evenson for a further discussion.)

¹¹ These export ratios are sensitive to patent law changes in other countries and require careful interpretation. Except for Japan, however, legal systems in most OECD countries are quite similar. Further, each country's system provides "national treatment" to foreign inventors (Evenson and Putnam).

Table 2. Inventions Used in Industries Supplying U.S. Agriculture

	Average Annual Inventions Patented							Import Ratio		Export Ratio
	1920s	1930s	1940s	1950s	1960s	1970s	1980s	1970s	1980s	1970s
311 Agricultural implements	198.5	154.1	115.4	179.7	186.7	201.4	170.8	.28	.61	.52
371 Agricultural chemicals	22.5	51.2	55.8	125.5	216.5	396.1	301.1	.52	.92	.76
3721 Chemical fertilizer	2.9	3.0	2.9	8.3	16.3	21.0	11.5	.39	.61	.99
3722 Mixed fertilizer	9.3	13.3	5.5	12.1	25.1	33.8	22.2	.28	.59	.51
3729 Other chemicals	10.3	32.9	57.4	105.1	175.1	341.3	268.4	.56	.97	.77
1053 Animal feed industries	25.7	29.6	19.8	30.3	39.3	58.3	41.6	.34	.54	.70
471 Grain elevator industries	3.6	2.6	1.5	2.7	12.9	5.9	5.4	.37	.44	.25
021 Services to livestock	212.1	152.2	34.4	171.1	185.8	218.3	222.1	.16	.31	.35
0211 Veterinary medicine	12.4	10.4	6.8	15.6	17.1	26.8	36.8	.28	.44	.54
0212 Animal breeding	1.4	1.2	1.0	1.6	2.3	2.5	2.7	.25	.41	.55
022 Services to crop product	218.0	198.0	124.6	232.5	227.0	228.2	192.8	.21	.37	.31
0221 Soil prep-cultivation	86.0	56.1	33.3	54.0	41.2	41.4	33.8	.19	.35	.25
0222 Dusting-spraying	10.4	11.6	7.2	13.9	14.2	14.8	18.2	.26	.42	.36
0223 Harvesting	67.3	66.1	46.9	107.7	109.4	105.0	80.8	.17	.35	.34
023 Services-managerial whole-saling services	10.0	8.1	5.8	9.4	12.4	19.2	18.1	.31	.51	.39
593 Seeds-chemicals	.4	.6	.5	.7	1.2	1.9	2.0	.18	.20	.43
571 Farm machinery	.5	.5	.5	.6	1.0	1.2	1.3	.33	.50	.44
501 Farm products	5.8	5.1	3.3	5.5	9.0	7.6	6.7	.18	.30	.29

ported (in parentheses). The first is the total number of domestically produced inventions of use in the four granting countries. The second is the ratio of total patents granted in the four countries to domestically produced inventions in the four countries. This ratio can be regarded to be an index of "transferability" for the technology in question. When this ratio is high, it in-

dicates that inventions originating in one country have a high degree of usefulness in another country.¹²

¹² These ratios have changed from the 1970s to the 1980s. The level of these indexes for 1970 is positively correlated with the change in the import ratios from the 1970s to the 1980s across industries. This suggests that the change in foreign competitiveness is higher in those industries with more inherent transferability.

Table 3. International Comparisons by Industrial Use Field

Granting Country	Trade Ratios		Inventions by Country of Origin						
	Import	Export	U.S.	U.K.	France	Germany	Japan	Canada	Others
A. Crop production (21,251, 1.94)									
U.S.	.35	.33	8544	421	270	470	446	29	1056
U.K.	1.68	.57	945	2256	317	892	196	31	1422
France	1.24	.26	913	352	3754	1384	258	17	1701
Germany	.64	.46	959	523	386	6697	231	13	2143
B. Livestock production (8,753, 1.57)									
U.S.	.30	.30	3541	152	99	234	146	25	405
U.K.	1.09	.44	373	1099	102	227	93	6	401
France	.81	.21	306	158	1642	361	83	6	423
Germany	.56	.33	367	172	142	2499	92	4	631
C. Services to farmers (14,918, 1.58)									
U.S.	.26	.33	6717	270	195	384	292	54	581
U.K.	1.26	.43	988	1863	150	416	164	16	605
France	.99	.24	570	257	2447	649	146	13	788
Germany	.55	.37	668	274	236	3897	193	13	765
D. Agricultural implements (6,219, 1.69)									
U.S.	.41	.47	2518	140	105	304	240	22	223
U.K.	1.43	.65	430	663	105	357	110	14	280
France	1.58	.41	332	129	852	478	81	3	330
Germany	.28	.66	432	159	145	2187	141	4	476
E. Agricultural chemicals (7,462, 2.39)									
U.S.	.66	.51	4011	501	212	789	508	5	636
U.K.	1.98	1.23	866	801	182	540	323	9	476
France	3.13	.71	733	325	477	478	279	8	520
Germany	1.25	.86	930	410	215	2116	382	7	695

The crop production farm sector received considerably more invention attention (2.4 times as much) than the livestock production sector. The services to farmer industries also received a high level of inventions, more than received by the implements and chemical inventions combined.

The transferability indexes indicate that inventions used in the agricultural chemicals industries are most transferable. The lowest levels of transferability (i.e., the most location-specific) are for inventions used in the livestock production industries and in services to farmers.

Table 3 reports import and export ratios for the four major inventing countries. Because the denominator of these ratios is the same, the differences in the ratio is a kind of net trade ratio (note imports from all countries are counted in the import ratio but only exports to the other three countries are counted). The United States, it may be noted, has a favorable ratio of exports to imports relative to other countries except in agricultural implements and to a lesser extent in agricultural chemicals, where Germany has a more favorable ratio.

The country of origin data show that the United States is the leading originator, i.e., domestic producer, of inventions in most industries and that Germany is the second leading originator of inventions. The United States is also the leading "consumer" or user of inventions in all industries.

Concluding Comments

The organization of invention data by IOU has only recently become feasible. This paper reports the first compilation of such data for the agricultural industries. The data series provide both historical and international insights. At this stage the data are intended to be informative indicators. It is expected that further refinement in the data series and methods may be achieved in the future. More important, the reporting of these indicators should provide scholars with inspiration for the further studies required to test hypotheses and to fully develop policy implications.

References

- Ellis, E. D. "The Philosophy, Construction and Uses of the Canadian Patent Data Base PATDAT" *World Patent Information*, 1981.
- Evenson, R. E. "Patent Data by Industry: Evidence for Invention Potential Exhaustion?" Mimeographed. New Haven CT: Yale Economic Growth Center, 1990.
- Evenson, R. E., and W. E. Huffman. "Science for Agriculture: The U.S. Experience." Yale University, 1990.
- Evenson, R. E., and Jonathan Putnam. "Institutional Change in Intellectual Property Rights." *Amer. J. Agr. Econ.* 69(1987):403-9.
- Kortum, S., and Putnam, J. "Estimating Patents by Industry: Part I." Yale University, 1989.

International Technology Transfer: Private Channels and Public Welfare

Margot Anderson and Bruce A. Larson

The economics of technology transfer is receiving increasing attention as economists acknowledge the role of transfer on future technological innovation, income, and growth. In agricultural economics, much research has focused on improving the international structure of intellectual property to facilitate technology creation and transfer or on analyzing the impacts of new technology, including that acquired from foreign sources, on productivity growth and welfare (Centner and White; Cochrane; Edwards and Freebairn; Alston, Edwards, and Freebairn). In general, this body of research is either not concerned with how technology is transferred or assumes the existence of a poorly defined technology market. But the specific channels of technology transfer are important in many sectors in the economy.

In agriculture, the increasing importance of the private sector in the creation and transfer of technology suggests a closer examination of the channels of transfer.¹ These channels include direct exports of goods that embody technology, foreign direct investment, technology licenses, and joint ventures. They can strongly affect the quantity of technology transferred, the distribution of gains to the agents involved, and any additional innovation in the importing country stimulated by the acquisition of the new technology (Cheng, Pugel). In addition, policies that either affect transfer directly (patent rights, licensing regulations) or indirectly (trade policies, industrial policy) can have a significant impact on the willingness of private companies to trade in technology.

The authors are economists with the Resources and Technology Division, Economic Research Service, U.S. Department of Agriculture. Senior authorship is shared.

The views developed in this paper are the authors' and do not necessarily represent the views or policies of the Economic Research Service.

¹ The private sector is a major source of new technology in agriculture. For example, it spent about \$2.3 billion in 1984 on research and development activities, accounting for about 60% of all such expenditures in agriculture (Pray and Neumeyer). Of this total, it is estimated that \$638 million was spent on chemicals, \$290 million on farm machinery and equipment, and \$954 million on human food (Pray and Neumeyer).

In typical transfers of technology, for example between input sectors, the technology supplier may be able to structure the contract for the transaction to increase the benefits received. The concept of a technology contract is more appropriate than price because technology transfer is often associated with financial payments, which may include lump sums and payments based on revenues or profits from the use of the technology, and nonfinancial conditions that specify how and where the technology may be used and in which markets goods produced using the technology can be sold. The technology supplier may also be able to include certain clauses into the contract to mitigate appropriability problems resulting from public good characteristics of the technology.

By altering the contract to fit the needs of the supplier and user, such channels may encourage technology transfer that would not occur otherwise in a market economy. For example, if market-like transfers are stifled because of market and/or government failure problems, these alternative transfer channels may also be welfare improving in a Pareto sense. How the gains from technology transfer are distributed across different groups will depend, therefore, on the attributes of the technology, industrial organization in the relevant markets, and the terms of the contract.

The structure of the contract for the transfer of technology is of primary interest here. To organize our discussion, the paper contains two main sections. First, following Rees, we develop an economic model of technology transfer using a principal-agent framework. This framework is applicable to a broad range of problems that seek to characterize the optimal form of a contract between two parties who face a range of actions and an uncertain state of nature. Looking at technology transfers as a principal-agent problem is useful because it allows us to analyze some important features of international trade in agricultural technology such as private sector transfers (and the payments associated with them) and uncertainty regarding agricultural and trade policies.

After the basic model and assumptions are outlined, several modifications to the basic assumptions are suggested to investigate how agricultural and property rights policies affect the incentives for technology transfer and the gains from such trade. Given the increasing role of the private sector in technology development and tensions in international policy arenas over trade barriers and internal supports, the economics of private-sector technology transfer is a crucial, yet poorly understood, component in assessing the effects of agricultural technology transfers.

Private Sector Technology Transfer

Consider a simple world with two countries, A and B, each with an agricultural output sector, an agricultural input sector, and a government sector. The agricultural output sector uses foreign and domestic inputs to produce goods for domestic and foreign consumption. Country A's agricultural inputs are also manufactured using technology developed in country B. Imported technology is acquired through two private channels: from one input sector to another, and from country B's input sector to country A's government sector.² Both governments also institute policies that directly and indirectly affect the development and transfer of agricultural technology. In terms of a principal-agent framework, the technology importing firm in country A is the agent, while the technology supplier in country B is the principal. The agent uses the technology and other inputs, a , to produce revenue, x . The particular value of revenue, x , depends on the agent's choice of a and the random variable, z , which the agent and the principal observe after a is chosen. The random component of the revenue function can be attributed to a number of sources. For our discussion, the random component represents shocks to the derived demand for agricultural inputs caused by uncertain agricultural and trade policies such as tariffs, quotas, and internal supports.³ Conse-

quently, the agent's revenue function depends on the degree and type of commercial policy used in the agricultural output sector.

In general, the revenue function of the agent who uses the principal's technology is $x(a, z; c)$ where c represents possible nonfinancial conditions or clauses specified in the contract for the use of the technology. For now, it is assumed that the agent and the principal know the revenue function and can observe the value of x . Let y denote the portion of the agent's revenues from the use of the technology that are used to pay the principal. The technology contract, defined as $Q(y, c)$, stipulates the terms under which the agent can use the principal's technology.

A basic issue revolves around whether the principal can observe *ex post* either the agent's actions a or the observed value of the random variable z . For the moment, we assume that the principal can observe either a or z , and through x can deduce the other. Thus, the basic model is one of Pareto-efficient risk sharing between the principal and the agent (Rees). Under these assumptions, the payment y depends explicitly on z , and the basic model involves the principal choosing $a, y(z)$, and c from relevant choice sets to

$$(1) \quad \max \int_0^1 u(y(z))f(z)dz \text{ s.t. } \int_0^1 v(x(a, z; c, z') - y(z), a)f(z)dz \geq v^*(z', z) \\ \text{where } v^*(z', z) = \int_0^1 v(z', z)f(z)dz,$$

where $u(y)$ and $v(x-y, a)$ are the principal's and the agent's Neumann-Morgenstern utility functions; $v^*(z', z)$ represents the agent's expected reservation utility; z' represents policies that affect the structure of intellectual property rights in the agent's country, such as patent protection and licensing restrictions; and $f(z)$ is the probability density function over the assumed range $[0, 1]$. The structure of international property rights, indicated by z' , enters the problem through the effect on the agent's revenues and reservation utility, v . This point will be discussed in the following section.

Following Rees, the variable a is assumed to yield disutility to the agent and is interpreted as the effort or expenditure by the agent in creating revenues. The choices of a and c are determined prior to observing the state of the random variable, but the payment for the contract may depend on the observed value of the random variable. Thus, in the risk-sharing case, the Pareto-efficient contract $Q(y^*(z), c^*)$ is the solution to the above problem (1).

² Public sector transfers are generated by the government sectors and include aid (capital and good transfers), exchange of scientific personnel, and international research activities.

³ This specification acknowledges an intertemporal component: input suppliers consider future demand for their products, which is directly affected by policies that influence agricultural production. The random component could also represent uncertainty associated with adopting new technology. For example, the technology may perform less well under certain environmental conditions, or certain preconditions that are not clearly known prior to using the technology must be met to use the technology successfully.

The model involves a true moral hazard problem when the principal cannot observe a and z . In this case, the technology contract must be based on observed variables to the principal, which is often assumed to be revenues x , so the contract in general becomes $Q(y(x), c)$. When moral hazard is present, an additional incentive constraint needs to be added to the model outlined in (1) because the principal can no longer design a contract to force the agent to choose the optimum defined for the Pareto-efficient risk-sharing solution.

Implications for Alternative Transfer Channels

While the above model may seem somewhat general, its basic implications can yield some insights into previous studies of agricultural technology transfer. For example, in an international context, it is not at all clear that the principal can observe the agent's actions a or the random variable z . One implication of such a situation is that the firm may have an incentive to change its basic structure to mitigate such problems. For example, rather than attempting to structure a complicated technology contract to deal with moral hazard problems, the principal could acquire an equity holding in the agent to improve observability of a , z , and/or x . In fact, the option of altering the basic structure of the firm, namely foreign direct investment, is one common form of international technology transfer and may be one method to mitigate moral hazard problems.

A transactions cost and monitoring story, which falls under the rubric of "internalization theory," is offered to explain foreign direct investment.⁴ In this view, direct investment involves a transfer of money capital and also a transfer of other resources including management services, technology, and "firm-specific knowledge" which, because of public good and monitoring problems, may be associated with appropriability problems. Thus, in this view, foreign direct investment may allow certain transactions to take place within multinational enterprises (MNE) that otherwise would not have occurred.

That foreign direct investment (FDI) is an important factor in international trade is clear. As

of 1984, the world stock of FDI was about \$599 billion, with the U.S. accounting for about 40% of the total. By 1987, U.S. FDI grew to \$309 billion (U.S. Department of Commerce). The United States is also the home for a substantial portion of FDI, about \$262 billion as of 1987. The quantity of trade destined for foreign-based subsidiaries, which is one indicator of the importance of foreign direct investment as a technology channel, is substantial in the U.S. economy and in the agricultural sector specifically. For example, in 1986, about 40% of all U.S. trade occurred between U.S.-based parent companies and their foreign affiliates, or U.S.-based affiliates of foreign based parent companies (U.S. Department of Commerce). In U.S. merchandise trade, exports from U.S. parent companies to their foreign based affiliates accounted for about \$57.5 billion, or about 27.5% of total U.S. merchandise exports (excluding mineral fuels) (U.S. Department of Commerce). As of 1986, U.S. direct foreign investment abroad totaled \$10.6 billion in food industries, \$23.0 billion in chemicals, and \$22.4 billion in machinery.⁵ In 1986, the foreign direct investment position in the U.S. was \$11.8 billion in food industries, \$22.9 billion in chemicals, \$3.8 billion in non-electrical machinery, and \$1.2 billion total in agriculture, forestry, and fishing.⁶

Once intrafirm trade is considered, the question of intrafirm prices, known as transfer and/or internal prices and the more general concept of a technology contract becomes important. For example, beyond direct pricing mechanisms, the U.S. Department of Commerce reports that intrafirm loans are a major source of funding for affiliated trade. The literature on multinational enterprises seems to view foreign direct investment as establishing similar objectives for the parent company (the principal) and the subsidiary (the agent), and that transfer prices can be used to the overall advantage of the MNE, perhaps at the expense of other countries, especially developing countries. As a result, many governments have established rules and regulations to limit the scope for transfer pricing.

Another important channel of technology transfer is technology licensing, which is a form

⁵ Of total U.S. MNE exports in agricultural-related industries in 1970, intrafirm trade accounted for 34% of the total in the food industry, 36% in chemicals, and 49% in farm machinery (Grimwade).

⁶ In the U.S. for 1977, related-party imports as a percentage of total imports were: 20% in beverages and tobacco, 23% in fish, 40% in fruits and vegetables, 19% in oilseeds, 38% in crude fertilizers and minerals, 43% in chemical elements, 23% in manufactured fertilizers, and 60% in machinery (Grimwade).

⁴ Foreign-direct investment occurs when a parent company establishes a subsidiary in another country or when a parent firm acquires a controlling interest in an overseas company, where a controlling interest implies at least 10% ownership of equity capital of the firm (U.S. Department of Commerce).

of nonequity international transfer of management services and technology that can occur between independent firms (and governments), a parent company and a subsidiary, or two firms involved in a joint venture. Licensing arrangements are contracts between an owner of technical knowledge and a potential user. The license contract often involves a lump-sum payment and a royalty fee based on a percentage of revenues or profits from the use of the technology. In 1981, for example, the United States was the destination of \$7.2 billion, or about 50% of total world licensing income, of which \$5.8 billion came from foreign affiliates (\$1.3 billion in developing countries) and \$1.4 billion was through unaffiliated trade (Grimwade).

Besides the payment stipulated in the licensing contract, clauses are often written into the contract that allow the principal to retain further control of the technology. Such clauses include restrictions on sharing information, using the technology for reasons other than that stipulated in the contract, and geographic location of production. In addition, certain standards of quality in the produced good may be stipulated, and firms may be required to purchase specific inputs in conjunction with the licensed technology.⁷

Katrak analyzes the case where a firm acquires a foreign technology through a licensing agreement that includes a fixed lump-sum payment and a payment based on a percentage of the firm's revenues. In the Katrak model, where both the technology supplier and the importing firm are profit maximizers and there is no uncertainty, it is shown that the importing firm prefers a lower payment based on revenues with a higher lump-sum payment for any total contract value. The logic is that a payment based on revenues reduces the incentives to innovate further after the transfer of the technology.

In terms of the model outlined in (1), when both the principal and the agent have constant absolute risk aversion r_a and r_p , respectively, a linear contract of the form $y(z) = bx(z, a, c) + d$ is an efficient way to share risk, where $b = r_a / (r_a + r_p)$, and d is a lump-sum payment. However, only when the agent is risk neutral is it efficient from the agent's perspective to push for a lump-sum technology contract because a lump-sum contract places all the risk of using the technology on the agent. Thus, there is a trade-off between risk sharing and the incentives for domestic innovation. Further analysis of this

trade-off is possible using the framework outlined in model (1).

Effects of Agricultural Policy on Technology Transfer

The pressure to restructure international agricultural policy is clearly evident. Since 1985, U.S. agricultural programs have reduced support levels and removed some of the incentives to increase yields. In the international arena, the current GATT negotiations are striving to reduce both internal and external support in developed countries in favor of a more market-oriented system. In addition, the developing countries are arguing in favor of reducing trade barriers on tropical products. While the induced-innovation hypothesis suggests that the development of new technology depends on domestic and international economic conditions, little attention has been directed to how agricultural policy influences the terms and, therefore, the welfare impacts of technology transfer.

Within the context of the model developed here, agricultural policies affect the agent's revenues and reservation utility level through the variable z .⁸ For a small country, which can take international prices as fixed, subsidies for producing commodities in the agent's country increase agricultural output and increase the derived demand for agricultural inputs, whether produced with the imported technology or not.⁹ Thus, the agent's revenues and reservation utility level may increase with a production subsidy. Whether the principal is also better off with the subsidy depends on whether the total value of the contract arises or falls with the policy change. For the case of constant absolute risk aversion discussed earlier, the variable portion of the payment will increase, but the fixed portion of the contract will rise or fall as necessary to maintain the agent at the new higher reservation utility level. Thus, in general, a production subsidy could improve or reduce the principal's utility.¹⁰ Conversely, removal or reduction of the production subsidy in a small economy case could result in a decline in commodity out-

⁸ Similarly, the agent's revenues and reservation utility can be determined partially by trade policies, such as tariffs or quotas on agricultural commodities. Any anticipated change in the level of a tariff can affect demand in the agricultural output sector for imported technology.

⁹ If there are quality differences between inputs produced with domestic and foreign technologies, then policy will affect the derived demand for the inputs differently.

¹⁰ Consumers are worse off in the agent's country if they finance the government, which incurs larger costs to finance the subsidy.

⁷ For example, for many years Argentine tractor producers who were subsidiaries of international firms were required to use local inputs in most phases of production.

put and reduce agricultural revenues. The agent's utility falls, but again the principal's utility may increase or decrease.

Currently, many developing countries face significant trade barriers on processed products through tariff escalation policies. Restrictions on exports of processed products has probably limited local technological innovation and dampened the demand for foreign processing technology. In terms of the model developed here, the agent's revenues from food-processing activities are reduced when trade policies restrict the demand for the agent's output. Removing or reducing trade barriers probably would increase the demand for foreign processing technology and/or increase the level of local research and development activities. However, as Mowery and Rosenberg point out, it may be difficult for a country to catch up after years of preferential treatment that allowed foreign producers to gain and maintain market share in processed products.

Restructuring trade policies may also influence the level of direct foreign investment in the agent's country, given that the proliferation of MNE's is partially attributable to tariffs and nontariff barriers that would prevent the principal from directly exporting to the agent. Reforming agricultural policies could increase export sales and remove some of the incentives for firms to invest in overseas subsidiaries, thereby having a dampening effect on technology transfer. For example, MNE activity is particularly strong in the pesticide and food-processing sectors, and a major contributing factor to increased overseas investment has been attributed to tariff (and nontariff) barriers to trade (UNIDO, Handy). The Argentine tractor industry provides an excellent example of the effect of tariff and agricultural policy on input industries. Until the late 1970s, Argentina's tractor industry was composed of subsidiaries of multinational tractor manufacturers. In 1978, a significant decline in MNE tractor production occurred due to policies that fixed agricultural prices below world market price and reduced tariff rates on imported tractors (Sievers).

Changing Intellectual Property Rights

There is a stylized view that improving the excludability of technology benefits, such as through changes in the international structure of intellectual property, will increase private-sector incentives to invest in technology develop-

ment, increase the transfer of technology across countries, and increase earlier disclosure of the innovation. The combination of these events is also thought to stimulate research by the firm (or country) that purchases the patented product.¹¹

In the model presented above, the agent's revenues and reservation level of utility depend on the degree of intellectual property rights protection in the agent's country, represented in (1) as z' . Within the current GATT negotiations, which contain provisions on international property rights, major technology-supplying countries are arguing for tighter restrictions on the use of foreign technology. The impact of a change in the structure of intellectual property rights can be examined using two simple scenarios. Let $z' = 0$ represent the case of no protection for foreign technology, and $z' = 1$ be the case of full protection. In the case of no protection, the agent uses the foreign technology without paying for the right to do so. In this case, the agent's utility obtained through pirating the technology, denoted here as $v^*(z' = 0, z)$, is greater than that which could be obtained from using some other existing technology, v .

If the structure of international intellectual property moves to full protection, $z' = 1$, the principal and the agent would have to agree on a mutually acceptable contract for the technology to be transferred. Because of the policy change, the agent's reservation utility level becomes $v < v^*(z' = 1, z) < v^*(z' = 0, z)$, and the principal can obtain some benefits from transferring the technology. Thus, the agent is worse off having to pay for the technology. The agricultural sector may be better or worse off, depending on the level of inputs produced with the legally imported technology; the quality of the input; and whether patent production encourages increased technology transfer from other suppliers. These issues are yet to be resolved.

Conclusion

Couching technology transfer within a principal agent framework allows us to focus on a number

¹¹ However, as Adikibi observes for the Nigerian case, the terms by which patents are actually utilized in the particular country determine whether in fact any knowledge is transferred/utilized in a country. In the Nigerian case, because of basic problems in design and implementation in the Nigerian patent systems as well as multinational corporation contracts for the economic exploitation of patented information, patents owned by foreign companies were not actually utilized in Nigeria. Thus, patents did not imply technology flow, disclosure of inventions, or further innovation based on the patented technology.

of salient features that describe international trade in agricultural technology. These features include the increasing importance of the private sector's role in transferring technology and uncertainty regarding agricultural and trade policies. An examination of the channels of transfer indicate different impacts on the transferring and receiving countries depending on the degree of uncertainty, the channel of transfer, and the type of public policies implemented. Assessing the costs and benefits associated with technology transfer will likely become more important in agriculture as countries face a more competitive environment in both technology and agricultural trade.

Further research needs to address these issues in more detail to help determine how contracting behavior changes over time or why some firms favor certain kinds of contracts depending on the type of technology transferred, the degree of uncertainty and the level of government intervention. In addition, we need to explore the implications of policies, other than reforms in property rights, that directly affect the creation and international flow of technology.

References

- Adikibi, O. T. "The Multinational Corporation and Monopoly of Patents in Nigeria." *World Develop.* 16(1988):511-26.
- Alston, J. M., G. W. Edwards, and J. W. Freebairn. "Market Distortions and Benefits from Research." *Amer. J. Agr. Econ.* 70(1988):281-88.
- Centner, T. J., and F. C. White. "Protecting Inventor's Intellectual Property Rights in Biotechnology." *West. J. Agr. Econ.* 14(1989):189-99.
- Cochrane, W. W. *Farm Prices: Myth and Reality*. Minneapolis: University of Minnesota Press, 1958.
- Cheng, L. K. "Optimal Trade and Technology Policies: Dynamic Linkages." *Int. Econ. Rev.* 29(1987):757-76.
- Edwards, G. W., and J. W. Freebairn. "The Gains from Research into Tradable Commodities." *Amer. J. Agr. Econ.* 66(1984):41-49.
- Grimwade, N. *International Trade: New Patterns of Trade, Production, and Investment*. London: Routledge, 1989.
- Handy, C. R. "The Globalization of Food Marketing." *Nat. Food Rev.* 13(1990).
- Katrack, H. "Payments for Imported Technologies, Market Rivalry and Adaptive Activity in the Newly Industrializing Countries." *J. Develop. Stud.* 25(1988):43-53.
- Mowery, D. C., and N. Rosenberg. *Technology and the Pursuit of Growth*. Cambridge: Cambridge University Press, 1989.
- Pray, C. E., and C. Neumeyer. "Trends and Composition of Private Food and Agricultural R&D Expenditures in the United States." *Agribus.* 6(1990):191-207.
- Pugel, T. A. "Endogenous Technological Change and International Technology Transfer in a Ricardian Trade Model." *J. Int. Econ.* 13(1982):321-35.
- Rees, R. "The Theory of Principal and Agent: Part I." *Bull. Econ. Res.* 37(1985):3-26.
- Sievers, M. "The Tractor Production of Selected Countries and the Transfer of Technology." Kiel, West Germany: Kieler Wissenschaftsverlag Vauk, Postfach, 1981.
- United Nations Industrial Development Organization. "Global Review of the Pesticide Industry Sub-Sector." Sectoral Work. Pap. Rome, 1988.
- U.S. Department of Commerce. *International Direct Investment: Global Trends and the U.S. Role*. Washington DC, 1988.

Technology Policy and Agriculture: Discussion

John Reilly and Roger Conway

The papers address two important forces affecting agricultural research and development: (a) how technology trade and innovation fit into an increasingly integrated world and agricultural economy and (b) the changing roles of public and private funding.

Evenson adds an empirical foundation to the issue of technology flows among countries, a major contribution despite the difficulties of using patent data as an indicator of inventiveness. Three broad points of interpretation are worth emphasizing. First, considerable caution should be used in interpreting a decline in U.S. patenting in the 1980s as a decline in inventiveness. Poor protection of intellectual property rights abroad and increasing internationalization may have provided a growing incentive for firms to use trade secrets even though property rights were protected in the major developed countries considered by Evenson. Comanor and Scherer, for example, found a secular decline in the propensity to patent. Additionally, changing quality of patents, relative propensity of large versus small firms or of firms in different countries to patent, differences in patent laws across time and countries, and relative likelihood of process versus product patenting all affect the accuracy of using patents to compare inventiveness across countries, industries, and over time. Bosworth suggested weighting a patent by the number of years it remains extant as a measure of quality. Scherer notes that larger firms tend to issue patents on minor advances at a great rate.

Second, if private sector inventiveness has actually decreased in the United States, it is likely indicative of reduced research and development (R&D) opportunity or effectiveness because, as

Frisvold reports, real private R&D expenditures in the United States have increased at an average rate of 2% per year during the 1970s and 1980s. This means the cost of maintaining agricultural productivity gains and the historical trend of declining agricultural commodity prices increased in the 1980s. The decline in inventiveness, however, is a big if, and the future implications are unclear. The 1980s may be one of the cyclical patenting downturns evident in Evenson's data. Such downturns may represent slow start-up periods during which research activity shifted to new opportunities. Biotechnology may have been such an opportunity in the 1980s with, perhaps, the patenting payoff yet to come.

Third, while the relative loss of U.S. inventiveness is consistent with the general observation that research has picked up abroad as many countries have caught up technologically with the United States, the welfare implications for the United States are unclear. In terms of overall national competitiveness, the important question is who appropriates the returns to invention. With internationally diversified stock ownership, multinational companies, and research and field-testing to take advantage of countries with favorable climates (in both the sense of weather and regulation/economics), the country where the invention is first patented may not be the country that appropriates the gain from a patent. Even accepting a relative loss of U.S. inventiveness, a broadening of inventive activity beyond the United States would be beneficial through lower world commodity or input prices. Another explanation of the trend is that gradual reduction in U.S. agricultural technology isolation may have led to an increase in technology imports indicating a shift in demand for imported inventions rather than a failure of U.S. inventiveness supply. U.S. companies may have, appropriately, shifted R&D funds from devel-

John Reilly is deputy director of the Resources and Technology Division, Economic Research Service, U.S. Department of Agriculture. Roger Conway is director, Office of Energy, U.S. Department of Agriculture.

opening new patented products to marketing already developed technologies. This interpretation would again lead one to welcome the trend toward increased technology imports because it would increase the competitiveness of the U.S. technology users.

Anderson and Larson analyze private firm, international technology transfer using a principal-agent approach. Focusing on the relationship between inventor and user raises directly the issue of how the gains to invention are shared in contrast to Evenson's data on who invents. The model highlights the need for firms to be flexible in exploring alternative strategies for transferring technology including, for example, arms-length licensing, creation of subsidiaries, and cooperative and joint development projects. It would be useful for the authors to develop a greater degree of structure to their model. While such structure would limit its applicability, greater insight would be provided into the specific conditions where it remained applicable. We raise several specific points below.

While there is a principal, i.e., the inventor, technology diffusion and licensing may include multiple agents. For example, an inventor may consider many agents until the agent (company or country) willing to negotiate the most favorable deal is found. In Anderson and Larson's model, the successful agent will be the one with the lowest level of risk aversion because this agent would be willing to accept the most risk. Large companies that can successfully pool risk across numerous new inventions may be willing to accept considerably riskier individual projects. If (companies in) developing countries are more risk averse, such a conclusion would explain a bias against rapid diffusion in developing countries. A single agent model also does not address why an inventor may license the invention to multiple agents rather than grant an exclusive license trusting the single agent to appropriate benefits across countries.

Anderson and Larson interpret the problem as one where the principal (the inventor) is relatively information-deficient relative to the agent. Presumably, the principal would have more complete access to the technology performance and therefore may have an information advantage. The agent, as well as the principal, could be the victim of moral hazard. The principal has the potential to greatly reduce the value of the licence by establishing additional licenses or by developing succeeding technologies.

The authors contrast their model with trans-

actions cost models, but the principal-agent formulation appears to be a special case of the transactions cost approach. Williamson identifies small numbers, information asymmetries, and bounded rationality as key elements that give rise to transactions costs and lead to internalization. The principal-agent problem is the special case where the small number is two, i.e., one supplier and one demander.

The authors also suggest analysis of the small country case where agricultural prices are fixed. One must exercise caution because, if the technology inventor fully exploits the technology, it means that the commodity price will fall in most cases.

The authors conclude that any restrictions on private sector transfers could result in less technology developed, as firms are constrained in earning the maximum return on their investment. The reduction in return may not be as great as one would initially imagine given that commodity trade could substitute for technology trade. If one envisions technology as a factor of production, then results of the Hecsher-Ohlin trade model yielding factor price equalization even with restricted trade in the technology factors apply, i.e., the technology would receive the same return whether traded or not if there were trade in commodities.

Frisvold, and Knudson and Pray deal with public versus private roles for R&D. Each provides a unique interpretation of the public R&D funding decision. A shortcoming of both papers is the treatment of the private R&D decision as exogenous. The more general model would view levels of each as jointly determined. In this regard, the Frisvold conclusion that the public sector should alter its research agenda to complement that of the private sector follows directly from the assumption that private sector R&D is predetermined. The apparent complementarity of some private and public R&D in Knudson and Pray's paper is a challenge to Frisvold's assumption of substitutability. Knudson and Pray's formulation of the SAESs as budget maximizers is an important and useful departure from the traditional assumption of fixed budgets. The results are intriguing but a fuller discussion of the data would be helpful. A more complete theoretical derivation of the empirical model would also help clarify the choices implicit in the model variants estimated. Because few of the coefficients remain significant and several change sign over the various specifications, it is difficult to draw general conclusions.

References

Bosworth, B. *Statistics of Technology, (Invention and Innovation)*. Review No. 28, Reviews of UK Statistical Sources. London: Heineman's Educational Press, 1980.

Comanor, W. S., and F. M. Scherer. "Patent Statistics as

a Measure of Technical Change." *J. Polit. Econ.* 77(1959):392-98.

Scherer, F. M. *Industrial Market Structure and Economic Performance*. New York: Houghton Mifflin Co., 1980.

Williamson, O. E. *Markets and Hierarchies: Analysis and Antitrust Implications*. New York: Free Press, 1975.

Technology Policy and Agriculture: Discussion

Marie E. Walsh

Agriculture is already technologically sophisticated and will become increasingly so as new biotechnologies and information technologies become available. Increasingly, however, agricultural technologies and practices are being questioned for their environmental, food safety, and animal welfare impacts. Society is concerned about the distributional impacts of new technologies as well as efficiency implications. Policy makers are faced with designing agricultural policies that consider these impacts. Agricultural economists can assist this process by evaluating the costs, benefits, and tradeoffs that society must face with respect to new technology development. They can play a potentially significant role in the design of technology policy to encourage the development of new technologies deemed most suitable by society.

The organizers of this session asked me to evaluate these papers from a policy perspective and to discuss their contribution to the policy-making process. As such, my comments are likely to be more critical of the papers and the agricultural economics profession than those of the other reviewers. In the area of agricultural technology policy, there is a genuine need to develop sound economic theory, to improve analytical methodologies, and to provide consistent data. In this context, all of the papers extend the knowledge base needed to analyze technology issues. All of the papers present tantalizing possibilities as to how they can be used in agricultural technology policy analysis. The Anderson and Larson paper provides a new way of thinking about international technology transfer and the roles played by intellectual property rights and agricultural policy. The Evenson paper verifies the growing realization that other countries have dramatically improved their technology development capabilities and provides a

much needed data set for further analysis. The Frisvold paper provides a framework for analyzing the allocation of public research funds to achieve market and nonmarket goals. The Knudson and Pray paper provides empirical analysis of the increasingly important issue of the privatization of public research. All of the papers provide improved theory, data, and analytical methodology needed to analyze technology issues, a necessary first step.

My criticisms of these papers and of agricultural technology papers in general arise from the frustrations of someone working in the agricultural technology policy field. Ideally, one would like to be able to find all of the information needed in the agricultural economics literature. However, rarely is the necessary analysis available. All too frequently, agricultural economics journals contain no articles relating to technology issues; and, when articles are present, they frequently are abstract, theoretical, and generic frameworks for evaluating an issue, much as the papers in this session are. Abstract frameworks do not present the type of information needed in a format that is useful to policy makers. Policy makers and their staffs have neither the time, resources, nor in many cases, the expertise to extend a theoretical paper to its logical conclusions and implications. In the arena of technology policy making, theoretical frameworks are often filed and ignored, and as such have a minimal impact on policy.

When empirical studies are conducted, they frequently fail to address the questions of policy interest. Efficiency and aggregate impacts are important, but they are not the only things that are important—issues of equity, distribution, intellectual property rights, environmental quality, food safety, animal welfare, and budget constraints are also on the minds of policy makers. They want to know how policies and technologies will differentially affect segments of the economy and how farms of different size and type and in different geographical regions will be affected.

The challenge for agricultural economists is

Marie E. Walsh is a policy analyst in the Food and Renewable Resources Program, U.S. Congress, Office of Technology Assessment.

The opinions expressed are those of the author and not necessarily those of the Office of Technology Assessment.

to provide policy makers with timely, relevant information on the costs and benefits of different technology policy options and to help them understand the tradeoffs that will occur. *Ex ante* analysis is generally preferable to *ex post* analysis. Rigorous quantitative analysis is desirable, but it is not the only way to participate in the agricultural technology policy debate. A well-constructed qualitative presentation of the pros and cons of specific policy alternatives and the relative direction and magnitude of expected impacts can be invaluable in the design of technology policy. Fortunately, *CHOICES* magazine provides a forum for agricultural economists to discuss important issues and to do so in English rather than mathematical equations, so non-economists can understand. Frankly, in the area of agricultural technology policy, articles in *CHOICES* are frequently more useful than those in the *AJAE* and other agricultural economic journals.

Evenson concluded his paper with the hope that the information provided will inspire scholars to perform further research to test hypotheses and develop policy implications. I wholeheartedly agree. My remarks have been critical on purpose. All of the authors have made an important first step. Sound theory is essential. But for policy analysis, it simply is not enough. I urge agricultural economists to do *ex ante* analysis on specific policy options and technologies and to broaden the issues that they address to include nonmarket values and distributional implications. Failure to do so will preclude agricultural economists from participating in the agricultural technology policy debate and from making contributions that they are so capable of making. Considering the increasing importance technology issues will have on agriculture, failure to participate actively in this debate is an embarrassment to the profession and one that I hope will not occur.

Technology Policy and Agriculture: Discussion

James F. Oehmke

The papers by Anderson and Larson, and Frisvold are similar in that each paper advocates the application of theories developed for other problems to technology generation and transfer: Anderson and Larson suggest greater application of the principal-agent paradigm, while Frisvold suggests greater use of industrial organization (among other) results. Although the advocacy of new applications is extremely ambitious, these papers make some progress toward their goals. The papers are similar in their style of argument. Each paper provides a brief description of the relevant theoretical results, discusses why some of the theoretical assumptions may (or may not) be appropriate for modeling technology generation or transfer, and gives a taste of policy implications that can be derived from these theories.

However, in part because of the ambitious goals, these papers are not completely successful. For example, Anderson and Larson's discussion of policy effects on the Argentine tractor industry is unconvincing: Why is a principal-agent framework needed to indicate that if prices are below the equilibrium level, then output will be below the equilibrium quantity? Similarly, Frisvold's discussion of quality-improving research does not seem to depend on the industrial organization of the market; demand shifts which raise equilibrium price tend to decrease the costs of a target price program when target prices are set above the equilibrium price under most imaginable industry structures.

Each of these papers would be greatly improved by focusing on one or two observations or empirical regularities which are not explained adequately by existing models of technology generation and transfer. This will entail greater discussion of the behavior to be explained and of the failure of existing models to explain this behavior. For example, Anderson and Larson would be more convincing if they compared the implications of the relevant principal-agent model

to other explanations of multinationals, such as noncompetitive markets (e.g., Krugman) or factor endowments (e.g., Dinopoulos, Oehmke, and Segerstrom), and found that the principal-agent framework best matched empirical observation. Perhaps the existing models do poorly in situations where risk is an important consideration.

Despite my reservations, I encourage the authors to pursue these lines of inquiry by developing more detailed models which yield testable implications. I agree with Frisvold's recommendation that revealed preference arguments may provide a better understanding of government behavior toward public research (e.g., Oehmke and Yao). I particularly agree with his recommendation for the establishment of a better data base on private R&D. It is impossible to determine which models have the most accurate behavioral implications without accurate data.

Evenson contributes to our understanding of private R&D by examining patent data. This paper does an excellent job of finding information on private R&D outputs, which is normally difficult to obtain, and of organizing the data in an informative manner. Three intriguing facts emerge: (a) the annual average number of U.S. agriculturally related inventions patented in the United States fell between the 1970s and 1980s, while the annual average number of foreign inventions patented in the United States rose over the same period (Evenson, tables 1-3); (b) the absolute number of U.S. patents has declined in the 1980s from the 1970s (Evenson, tables 1-3); and (c) the United States has been a net exporter of patentable, agricultural inventions to France, Germany, and the United Kingdom over the 1960-87 period (Evenson, table 4).

The data presented in this paper raise a number of important questions: Does the recent decline in the number of patents portend a time of stagnating agricultural productivity? Is the increased percentage of U.S. patents originating in foreign countries a result of technological factors such as a relative decline in U.S. research productivity, monetary factors such as fluctuating exchange rates and the strong U.S.

James F. Oehmke is an assistant professor, Department of Agricultural Economics, Michigan State University.

dollar, or other causes? Is the increased importation of foreign technology a signal of declining U.S. technical competitiveness, or is it socially beneficial because it makes improved technology more easily available to farmers and processors?

Pray and Knudson focus on the interaction between the public and private research sectors in a political economy framework. As Frisvold points out, this is an important direction of research. Pray and Knudson's regression analysis shows the importance of the private sector and incentives designed for the private sector in determining the level of public research expenditures. The three most interesting explanatory variables in this analysis are private research expenditures, which have a negative but not statistically significant effect on public expenditures; potential public-sector royalties from public-sector research, which has a positive and in one case statistically significant effect; and industry contributions to experiment stations, which has a positive and significant effect. Because the data are separated by commodity, the last result can be interpreted to mean that private contributions can redirect the commodity focus of experiment station research.

The regression analysis could be improved in further work. Annual observations should be available for the data. Use of annual data would allow investigation of questions such as incrementalism in the budgeting process. The number of observations would be increased sufficiently to test hypotheses about differential effects of private sector actions on different commod-

ities. While such hypotheses are discussed in the text, they are not testable in the current regression framework.

One peculiarity in the paper is the discussion of private sector "crowding-out" of public sector research. This direction of effect suggests that the public sector has a comparative advantage in research relative to the private sector. While those of us committed to public institutions accept this as fact, the industry as a whole may not be so lenient.

Finally, Pray and Knudson examine the implications of their regression analysis for optimal allocation of research expenditures. Using their (admittedly crude) measure of social benefits, they find that industry contributions to the public sector push the cross-commodity allocation of public research funds closer toward a social optimum. This intriguing phenomenon deserves more thorough empirical investigation and conceptual explanation.

References

- Dinopoulos, E., J. F. Oehmke, and P. Segerstrom. "High-Technology-Industry Trade and Commercial Policy." Dep. Agr. Econ. Staff Pap. No. 89-43, Michigan State University, 1988.
- Krugman, Paul R. "Increasing Returns, Monopolistic Competition, and International Trade. *J. Int. Econ.* 9(1979):469-79.
- Oehmke, J. F., and X. Yao. "A Political Preference Function for Government Intervention in the U.S. Wheat Market." *Amer. J. Agr. Econ.* 72(1990):631-40.

The Effect of the 1990 Farm Bill on Agricultural Trade

George E. Rossmiller and Rachel A. Nugent

A struggle has taken place across the Atlantic between the pocketbooks of the U.S. federal government and the European Economic Community during the 1980s. The battlegrounds have been importing countries that have received heavily subsidized exports of agricultural goods from the two big producers. The casualties are competing exporters and those, often third world countries, who would produce and perhaps even export agricultural products in an undistorted world market. Two major policy events during 1990 threatened to alter the weapons and fortunes in this trans-Atlantic struggle: reauthorization of the five-year U.S. farm legislation and the global trade negotiations under the GATT. This paper reviews briefly the status of these two events and summarizes expected trade impacts.

Both events promised to address long-standing imbalances in agricultural markets. The combination of policies in the United States and EC, increases in farm yields, and saturation of domestic markets had led to recurring conditions of excess capacity, turbulent farm sector economies, and competition for international markets. These trends began with policy changes in the early 1970s that contributed to huge government outlays as the 1980s progressed.

Pressure to alter the policy course became severe in the United States as government support for agriculture soared and high loan rates in the 1981 Farm Act curtailed U.S. markets overseas. The 1985 U.S. farm legislation increased market competitiveness by lowering loan rates and maintained target prices at 1985 levels for two years before phasing in a 10% reduction between 1988 and 1990. The EC was likewise making marginal policy changes that would reduce its budgetary exposure and lower payments to farmers on selected commodities. The result was eventual lowering of budget outlays, but not

before they peaked at an all time high of \$26 billion in the United States and over \$30 billion in the EC in 1986.

Countering the trend toward lower support, both the EC and United States were actively vying for export markets by selling commodities at substantially cut-rate prices. The United States initiated the Export Enhancement Program (EEP) to directly counter the EC's export subsidies in designated markets.

The strong budgetary commitment continued steadily until a pause caused by the North American drought of 1988. World market prices rose as a result, and U.S. and EC government costs of agricultural programs plummeted. Strengthened prices and lower subsidies did not last, however, as the underlying trends causing surpluses resumed.

1990 Farm Bill

The debate over the 1990 U.S. farm bill began as a whisper. On most issues, participants were content to maintain the elements of the 1985 farm bill, at least until the outcome of GATT trade talks was known. It was widely expected that the real changes would come after a GATT agreement and would be incorporated into a 1991 farm bill.

The 1990 farm bill did make some change in farm programs. The legislation written during the summer and fall froze target prices throughout the life of the bill, controlled production through acreage reduction and conservation reserve programs, and increased farmers' flexibility in planting decisions. The bill carried a pricetag of about \$55 billion for the 1991-95 period—about \$2 billion-\$5 billion more than continuing the features of the 1985 farm legislation over this period.

In October, when the drafted farm legislation was being finalized by a conference committee, Hill negotiators and the Bush administration

George E. Rossmiller is director, National Center for Food and Agricultural Policy, Resources for the Future; Rachel A. Nugent is an economist with the National Security and International Affairs Division of the General Accounting Office.

agreed on a deficit reduction package to avoid the more stringent budget cuts that would have been necessary under the Gramm-Rudman-Hollings Budget Reduction Act. Farm support programs were among the most severely affected. Budget negotiators informed the agriculture committees they must cut \$13.6 billion from the five-year legislation—about a 25% decrease—with \$1.4 billion in immediate cuts for fiscal year 1991. This was still less than the 32% cut that would have been required under Gramm-Rudman-Hollings for fiscal year 1991. Direct support reduction in 1991 is estimated at about \$500 million under the Budget Reconciliation Act versus an estimated \$2 billion reduction had a Gramm-Rudman-Hollings sequestration taken effect. The budget summitters effectively forced some policy changes in farm programs that were politically infeasible if left to the agriculture committees on the Hill.

One of the primary cost-cutting measures adopted was the so-called "triple-base" for cereals and cotton. The triple base reduces crop acreage eligible for support payments by a mandatory 15% (and an additional optional 10%) while allowing farmers to grow crops on that acreage (including program crops) and sell those crops at whatever price the market will provide. In effect, a farm is divided into one section that is eligible for full program payments, another section that must be idled to comply with acreage reduction requirements, and a third section that can be planted in any crop (other than fruits or vegetables) to provide earnings not subject to government subsidies. The triple base has the effect of decoupling production at the margin from government payments. This will likely prove to be the most important policy change in the 1990 legislation.

Some of the other notable budget-cutting measures included in the deficit package are a change in the calculation of average market prices from a five-month to a twelve-month basis so that a higher market price basis is used to calculate deficiency payments, new loan origination fees and service charges for program participation, and a move to guaranteed loans from direct loans for Farmers Home Administration and Rural Electrification Administration lending programs.

The budget reconciliation bill contains some major adjustments in the event there is no GATT agreement by July 1992. Specifically, \$1 billion must be provided for export promotion (EEP), a marketing loan program for wheat and feed grains must be instituted, and acreage reduction

minimum requirements may be waived. If an agreement is concluded but not implemented by July 1993, the secretary of agriculture must consider waiving the spending reductions in the bill, increasing funding for export promotion, and instituting a marketing loan for wheat and feed grains. This clause was insisted upon by Capitol Hill in an effort to put additional pressure on the European Community in the GATT negotiations.

Impact on Trade

The effect of the combined 1990 farm bill and the budget reconciliation act on the composition and level of agricultural production and trade is unclear for several reasons. First, the triple base allows farmers to plant alternative crops on up to 25% of their program crop base without losing program base acres. Presumably, farmers will make this planting decision, at least at the margin, based on expected market returns of the commodities they are technically able to produce.

But government support for farm incomes is substantially reduced. With the frozen base yield now significantly below trend yield and the 15% mandatory reduction in program crop acreage eligible for support payments, fully 26% of corn production and 23% of wheat production in 1990 would have been ineligible for support. The percentage of production ineligible for support can be expected to increase with rising yield trends but may vary among commodities depending on shifts in acreage planted to the different crops.

Second, the 1985 act had introduced environmental and conservation compliance requirements that are viewed by many producers as punitive. The 1990 act further restricts eligibility for federal farm programs when producers are not in compliance with their conservation plans. Third, while the \$50,000 deficiency payment limitation per person was retained, the aggregate payment limit was halved from \$500,000 to \$250,000.

In sum, it is likely that program participation will be discouraged by the reduced acreage eligible for support payments afforded by the combination of the triple base and the acreage reduction program, more stringent rules and penalties on environment and conservation compliance, lower payment limits, and lower deficiency payments derived from using a twelve- rather than a five-month average of market prices. Lower program participation is likely over the

life of the 1990 act, particularly in years of expected strong market prices.

In 1991, however, participation rates are likely to remain at or above present levels. This can be attributed to the tumble in commodity prices in recent months, particularly since mid-1990. For example, November 1990 wheat prices were down 31% from November 1989 and 24% from the second-quarter 1990 average. Corn prices in November 1990 were 8% lower than in November 1989, but 16.7% lower than their second quarter 1990 average. Comparable rice price percentages were 12.8% and 4.8%. Cotton prices have been stable during the past year showing a 1.1% increase between November 1989 and November 1990. Soybean prices were down 2.8% from November 1989 to November 1990, with meal prices down 10.6% and oil prices up 5.2%. Thus, with lower prices one might expect participation rates to remain at present levels or even increase until markets recover.

These price relationships also provide clues as to the likely shifts in the production mix, at least for next year, under the triple base. With wheat prices the most depressed of the commodity prices, it is likely the wheat triple base acreage will be shifted to other crops where possible, toward cotton in parts of the southern plains, toward corn in the fringes of the Corn Belt, and perhaps to minor oilseeds in areas where the market channels and processing facilities are available.

The net result of various responses to the triple base is expected to be eventually lower participation rates (after depressed market prices recover), therefore increasing production. The U.S. output and its mix will become more responsive to world market prices. Yet, world market prices may decline with overall increases in U.S. production and exports.

In addition to the triple base, several new program changes may affect trade quantities. Increasing loan rates to 85% of a five-year moving-average market price, eliminating the high and low prices, instead of 75%, ostensibly raises the price floor for wheat and feed grains. However, several complicated adjustments are required or allowed to retain U.S. competitiveness in international markets. A mandatory downward adjustment in the loan rate of zero to 10% depending on the stocks-to-use ratio and a discretionary 10% downward adjustment by the secretary of agriculture to maintain export competitiveness are included in the act. A new marketing loan for oilseeds and limits or potential removal of acreage set-asides may benefit ex-

ports through lower market prices and increase production.

Uruguay Round Effects

The Uruguay Round of international trade talks was expected to bring about a brave new world of global trade relations. Our sights at the moment are considerably lower. In fact, the question of whether an agreement will be reached in agriculture is still open at this writing. The issue has boiled down to the EC interest in reducing direct budget costs of domestic farm programs by reducing support levels but without fundamental reform of the Common Agricultural Policy and erosion of community preference, against the U.S. and Cairns Group interests in increasing market access and placing limits on the ability of all countries, including the EC, to subsidize exports. Because the U.S. and the Cairns Group positions would require fundamental CAP reform that would put community preference in jeopardy, a substantial GATT agreement is not likely in the cards this time around. Thus, it is not likely to have a major impact in liberalizing trade or reducing trade tensions.

Part of the reason the Europeans have offered so little of substance in their negotiating package is the preoccupation in Europe with EC-92 market harmonization, the breakdown of the command economies and the struggle to replace them with market economies in Eastern Europe during the past year, and German reunification. There are substantial potential costs facing Western Europe in handling these events, and even greater costs if they do not succeed in dealing with those issues. In contrast, it appears that a GATT agreement would cost great political and economic stress in Western Europe in the short run, while an agreement presents only vague and intangible long-run benefits.

The EC is likely to become more inward-looking whether or not a GATT agreement is reached. Already heavily focused on the formation of EC-92 institutions, and in need of markets for excess production, the EC is likely to expand membership to include nations from both Western and Eastern Europe. If this significantly expanded trading bloc were to be highly protectionist *vis-à-vis* the outside world, the trend toward formation of trading blocs by other countries would be exacerbated.

This scenario implies a world in which bilateral agreements and independent trading blocs may become the dominant trade structures, thus

further weakening the GATT. Agriculture in the United States and EC would remain heavily subsidized. Developing countries in Africa and parts of Latin America would be kept out of the bidding, with no opportunity to establish growth patterns based on international competitiveness.

Finally, it is accepted wisdom by now that external macroeconomic factors affect U.S. farm income and stability to a far greater extent than previously believed. The world market takes roughly one-fourth of U.S. agricultural output. Changes in demand caused by developing country debt problems, reconstruction in East Europe, oil price shocks, and other macroeconomic impacts greatly alter the potential health

and trade prospects of the U.S. farm sector—more so, in fact, than the 1990 farm bill.

References

- International Monetary Fund, Commodities Division of the Research Department. Monthly report. Washington DC, Dec. 1990.
- U.S. Department of Agriculture. *The 1990 Farm Act and the 1990 Budget Reconciliation Act*. Washington DC, Nov. 1990.
- U.S. Department of Agriculture, Economic Research Service. PS&D Time Series, annual data. Washington DC, 1990.

The 1990 Farm Bill and the Uruguay Round

C. Ford Runge

When former Secretary of Agriculture Clayton Yeutter accepted the position, two clear priorities for his term were evident. The first, domestic, priority was to move as far as possible in the direction of a safety-net program for farm income support in the 1990 farm bill. The second, trade-related, priority was to put together an acceptable package in the Uruguay Round of multilateral trade negotiations, in which agriculture would play a central role. This short paper reviews the progress of both of these efforts to date, and assesses the impacts of this progress on the commodity composition, income, and asset values of American agriculture.

This assessment is qualitative. The detailed numerical estimation of these effects is left to those equipped with the large models necessary to estimate them. The assessment is also affected by a variety of external factors which loom especially large at the beginning of 1991. Apart from the unsettled outcome of the Uruguay Round and the economic recession, these include the political and military situation in the Middle East and the breakdown of functioning economies in Eastern Europe and the Soviet Union. Together, these external forces are likely to play a much larger role in determining the fortunes of American agriculture than either the 1990 farm bill or the Uruguay round, at least in the short term.

The 1990 Farm Bill

The farm bill debate of 1990 occurred in essentially two phases. The first was an exercise in political fantasy in the spring and early summer, in which the members of Congress indulged the wishes of a variety of commodity groups with promises of increased levels of support. These

domestic promises were linked to a variety of stern statements that the General Agreement on Tariffs and Trade (GATT) would not lead the farm bill to be "written in Geneva." Fears over this (unlikely) prospect were part of the reason that the 1990 bill ultimately contained "snap-back" provisions which guaranteed that if the Uruguay Round "failed," farmers would be compensated with extra spending on the Export Enhancement Program (EEP), and marketing loans would be instituted for coarse grains and wheat.

During the fantasy phase, the administration issued its own version of a wish list, a green-colored document detailing its proposals for a whole farm base, or Normal Crop Acreage (NCA) scheme, together with a variety of other more-than-incremental proposals, such as the elimination of crop insurance in its current form. While not quite as "dead on arrival" as the Reagan administration's 1985 proposal, the fact that the proposal contained no budget numbers indicated that the administration was engaged in a strategic bargaining exercise.

The second phase of the farm bill process was driven by budget realities, during which most of the fantasies (though not all) were laid to rest both in Congress and at the U.S. Department of Agriculture (USDA). As the "budget summitteers" flailed away in attempts to conform to the Gramm-Rudman deficit reduction targets, it became evident that even major attempts to staunch the flow of red would not contain the hemorrhaging federal budget, especially as recession deepened. Agriculture spending, however, was a virtually unanimous candidate for cuts, and as the need to eliminate roughly \$13.6 billion over five years emerged from the budget talks, it dictated certain moves in the agriculture committees and at USDA that saved money while saving seats.

Because target prices are visible political numbers and had been the focus of much of the fantastic promises made by incumbents seeking re-election, the least politically damaging way

C. Ford Runge is an associate professor and a member of the Center for International Food and Agricultural Policy, University of Minnesota.

Financial support from the Northwest Area Foundation, St. Paul, Minnesota, is gratefully acknowledged.

to find budget savings was by reducing the number of base acres in the major budget programs (coarse grains and wheat) that were eligible for deficiency payments. This budget pressure dovetailed (though not perfectly) with the idea of "flexibility" which had underpinned the administration's argument for NCA. Because total flexibility under NCA was neither necessary to achieve the budget targets nor desirable to many commodity groups and their supporters in Congress, a "triple-base" emerged as a natural compromise. While much else in addition was done or undone in the 1990 farm bill, the "triple-base" was at its heart and was driven primarily by budget pressures rather than deeply felt attraction for a clear step in the direction of decoupling. Such a step it was, however, giving the administration a relatively strong hand going into the December 1990 GATT meetings in Brussels, when it could claim virtue for having moved in the direction of its proposed safety net.

The impacts of the 1990 bill on domestic producers will be mixed. One notable aspect of the "triple-base" is that it plays much more to the advantage of the corn/soybean grower than the wheat grower, especially in the North and Northwest where flexibility, apart from wheat and barley, means canola, sunflowers, or dry edible beans, and not much else. Because neither fruits nor vegetables are allowed on "flex-acres," in essence the bill provides for more oilseeds in lieu of either corn or wheat. In total, the acreage shifts are not likely to be substantial unless the economic environment becomes favorable enough for large numbers of producers to abandon the programs altogether and grow what they please. In the short term, however, the continuing secular decline in wheat prices, aided and abetted by the interaction of EEP and the European Community's (EC) export restitutions, will make farming hard in areas like the Palouse. One curious consequence of the triple base was to promote opposition on the part of many wheat growers to increased acreage reduction programs (ARPs). Many of the same growers had argued in 1985 for mandatory production controls.

While a paper of this length precludes commodity-by-commodity analysis, a word about the dairy sector is in order.¹ Both the 1990 farm bill

and the budget reconciliation act contain provisions which are significant for the dairy sector. Support prices are not to fall below \$10.10 per hundredweight (for 3.67% milk) for several years. And the producer assessment, while only 5¢ per hundredweight the first year, could rise substantially for producers that expand production. This latter provision provides a mechanism for extracting payments in the event supply exceeds demand and may be exercised if output continues to drift upward.

If the experience of the 1990 bill does not discredit "flexibility," even greater steps can be taken in the future to decouple payments from production, to reduce ARPs, and to loosen requirements forbidding non-program crops from being grown, so that the real advantages of NCA can be realized. One area where the 1990 bill fulfilled growers' fantasies beyond all expectations was the sugar regime, which, despite a negative GATT panel ruling, emerged not only unscathed but arguably enriched from the legislative process. This should tell the analyst something about whether farm bills are written in Geneva. Soybeans, which had been held up for years as a model of "market orientation," threw in the towel and sought the protection of a marketing loan, which is in essence no different from the EC's restitutions, although it is set in such a way as to do the soybean grower little good. Having compromised on principle, the soybean growers failed to bring home much of a prize.

The "snapback" provisions of the bill allow for both more EEP spending and the extension of marketing loans to coarse grains and wheat if GATT "fails." These provisions contain the worst elements of both fantasy and reality. On the one hand, they are unlikely to be funded at levels which would realistically be required to truly punish the EC for its intransigence unless the \$13.6 billion spending target is abandoned. On the other hand, they will surely provoke retaliation, in all likelihood leading to even lower prices, especially in the wheat market. It is the threat of such retaliation that would make failure in the Uruguay Round of real concern.

The Uruguay Round

As of this writing, the likelihood for a meaningful package of reforms in the areas of marketing access, internal supports, and export subsidies in GATT are slim. When the EC, together

¹ An excellent briefing on a commodity-by-commodity basis is "The 1990 Farm Act and the 1990 Budget Reconciliation Act: How Farm Policy Mechanisms Will Work Under the New Legislation." Washington DC: U.S. Department of Agriculture, Nov. 1990.

with Japan and South Korea, rejected the Hellstrom compromise proposal in Brussels on 6 December, they signaled that, even if a final deal is achieved, it will fall short of the proposed compromise. That compromise called for 30% reductions in both export subsidies and internal supports on a base year of 1990 (as distinct from the EC's proposed base of 1986) and 30% increases in market access over five years, with a minimum 5% market access guarantee at the outset.

The prospects in GATT are either for something short of the Hellstrom compromise, or nothing at all. In terms of immediate impact on the U.S. farm sector and its balance sheet, either outcome would take several years to show up, unless a trade war erupted quickly in the face of failure. Something close to the Hellstrom compromise would reinforce the logic the "triple-base" by mandating further reductions in deficiency payments and would create an excuse for ending the ill-advised EEP program. The market access provisions would also assist in lower import quotas in sugar and to a lesser degree in peanuts, dairy, and other border-protected commodities. These effects would occur over a relatively long time (5–10 years), giving the farm sector plenty of opportunity to adjust.

One of the ironies of the debate over the Uruguay Round in farm circles has been the paranoia GATT has produced, which has been fed by neopopulist opponents of liberalization. These opponents are usually admirers of supply control and sometimes of the European Community. If the GATT talks fail completely, the irony will be that the retaliation mandated by the "snapback" provisions of the 1990 bill will actually fan the flames of protectionism, leading to attacks on the EC, Japan, and South Korea, which together constitute huge agricultural export customers. If farmers are looking for something to be paranoid about, it should be a trade war, rather than GATT. Such a trade war will have two primary effects. First, it will further depress world markets, leading to even lower commodities prices, especially in the wheat market. Second, it will cost money, which, unless Congress is prepared to reverse its stand on agricultural spending, could mean even less for deficiency payments. If the trade war spreads beyond agriculture to include other sectors of the economy, it would deepen the current global recession, lowering profits and government revenues, putting even more downward pressure on both the demand for agricultural exports and the

ability of government to subsidize them and the farm sector.

Conclusion

What are the overall prospects, given the 1990 farm bill and the unsettled trade picture, for U.S. agricultural incomes, land values, and commodities prices? Essentially, the 1990 bill has set a course leading to greater market orientation, although a few exceptions remain, such as sugar. Yet, the trade picture may mean that at the border, and in the world market, this market orientation is frustrated by the protectionist elements of the world community, notably the EC.

Whether or not a trade war erupts, I would expect the level of farm incomes to stay at roughly current levels or to fall if recession cuts into global demand. The commodity mix of U.S. agriculture will not be affected greatly by the triple-base unless larger numbers of farmers leave the program, which is unlikely in a recession. The linkage from export demand to farm incomes and in turn to land values (see Runge and Halbach) implies that land values will not rise appreciably either and could fall if a trade war erupts.

However, several factors are still unknown which could radically change this picture, probably for the worse. One is the prospect of continued instability in the Middle East, which could raise the price of petroleum and its derivatives and thus farm input prices far more rapidly than it would boost commodities prices. The second is the collapse of the Eastern European economy, which, while apparently offering a market for U.S. grain, constitutes less a paying market with hard currency to spend than a charity case. Both events likely would deepen global recession. In light of the current popularity of bashing our allies for failing to support our efforts in the Middle East, it would not be surprising to see the bashers join forces with the protectionists (many are already one and the same) to call for retaliation against these same parties in agriculture. It is far easier to blame lost jobs or farm foreclosures on Japan, South Korea, or the EC than it is to acknowledge that recession breeds protection, which deepens recession further. This was the same cycle so clearly exposed by Keynes in *The Economic Consequences of the Peace*.

This discussion began by noting that Clayton Yeutter stepped into his position at the Department of Agriculture with at least two broad goals. His efforts to bring greater market orientation to

agriculture must on balance be judged a success, although they are a patchwork, with success in the "triple-base" offset by persistent distortions in areas such as sugar and dairy. His other and related objective—to begin the process of agricultural trade liberalization in the Uruguay Round—foundered on the resistance of the EC, coupled with global recession, tension in the Middle East, and the collapse of Eastern Europe, none of which were entirely predictable. These forces may alter the course of history in ways inimical to greater trade liberalization, although I hope that the leaders of the free world will not err today as they did in

the 1930s, with respect to either protectionism or appeasement.

References

- Keynes, J. M. *The Economic Consequences of the Peace*. London: Macmillan and Co., 1919.
- Runge, C. F., and D. W. Halbach. "Export Demand, U.S. Farm income, and Land Prices: 1949–1985." *Land Econ.* 66(1990).
- U.S. Department of Agriculture. "The 1990 Farm Act and the 1990 Budget Reconciliation Act: How Farm Policy Mechanisms Will Work Under the New Legislation." Washington DC, Nov. 1990.

Impacts of the 1990 Farm Bill on Consumers

Carol S. Kramer

Any U.S. farm bill, including the 1990 "Food, Agriculture, Trade, and Conservation Act," affects domestic and foreign food consumers in numerous important ways: through potential impacts on food availability, food costs, food safety and quality, marketing rules, nutrition provisions, and overall research agenda. In addition, the act influences consumers of environmental amenities explicitly through environmental provisions and indirectly through commodity and other titles. Finally, the farm bill costs money, which consumers pay as taxpayers. This paper highlights some of the provisions of the 1990 farm, appropriations, and budget reconciliation bills likely to prove most significant for consumers in the short and longer term. It also discusses briefly the status of trade negotiations pertaining to "harmonizing" food safety standards for traded goods. Finally, the point is made that much significant food-related policy is not included in the farm bill but has been passed or can be expected to pass piecemeal in other forms of legislation.

The "Food, Agriculture, Trade, and Conservation" Act of 1990

Washington-based analysis of the 1990 farm bill has vacillated between depicting it as largely business-as-usual and a disappointment to reformists, on the one hand; and, on the other, as a dramatic step toward greater market orientation and planting flexibility, with some important new consumer and environmental reforms. The truth, as usual, resides between.

The former view traces the roots of the new act to the Farm Security Act of 1985, as influenced by federal deficit and budget contingencies. It points to the little-changed peanut, sugar, cotton, tobacco, and dairy programs and the newly instituted soybean loan rates. The reformist view emphasizes the triple-base provisions

of the bill, which remove 15% of base acreage from eligibility for commodity payments but permit production of some alternative crops. New environmental and consumer measures include pesticide record-keeping requirements, expanded environmental compliance measures, and certain changes in marketing rules or research affecting fruits and vegetables, grain quality, and organic products.

Complicating assessment of potential farm bill impacts is the massive uncertainty in the global political economy. Uncertainty stems from events in the Persian Gulf, the USSR, the stalled GATT negotiations, German reunification, Eastern Europe, global recession, the vulnerability of the U.S. banking system, among others. At issue is whether the incremental policy impacts of the farm bill will be swamped by the macroeconomic tides loosed by these major political events.

Given all the uncertainties, what does the new act mean for consumers? One could argue that for the consumer the 1990 farm bill may mark a new watershed in some areas, but one that is so far only dimly visible to the careful observer. Principally important are the "green" measures and tendencies in the farm bill: requirements to develop national organic certification standards, to require farm pesticide recordkeeping, to establish new water quality measures, to undertake research into the effects on pesticide use and consumer demand of "cosmetic" grade and quality standards for fruits and vegetables, and to redirect substantial research expenditures toward food safety, sustainable agriculture, and environmental improvement. In addition, in a time of budget austerity, the food assistance programs are reauthorized without significant change: without cutbacks, and containing some measures to make them more accessible to eligible recipients. Finally, outside the farm bill, important new nutrition labeling and nutrition monitoring legislation passed.

The "Food, Agriculture, Trade, and Conservation Act" of 1990 (FATCA) comprises twenty-five titles in a five-year framework for agricul-

Carol S. Kramer is a fellow, National Center for Food and Agricultural Policy, Resources for the Future.

tural and food policy; the first eleven titles deal with specific commodity programs and general commodity provisions. A subsequent mix of titles includes such areas as forestry; conservation; agricultural trade; credit; research; rural development; food assistance; fruits, vegetables, and marketing; grain quality; and organic certification.

The 1990 farm bill has been widely touted for the future cuts implied for farm spending. Development of the farm bill must mesh with the agricultural appropriations process and with the overall budget process as well. Farm appropriations bills begin in the House and then proceed to the Senate, after which a conference committee resolves the differences in light of budget requirements. As agriculture's share of a five-year \$500 billion deficit reduction package, farm spending is now slated at \$40.4 billion, down 25% from the earlier Office of Management and Budget baseline projections of \$54 billion. Of the \$13.6 billion cut during the budget summit, approximately \$8 billion reflects the 15% reduction in cropland acres eligible for farm subsidy payments. Other major cuts are in direct loans made by the Farmers Home Administration and the Rural Electrification Administration. Costs of domestic food assistance entitlement programs reauthorized under FATCA are not expected to change substantially as a result of the bill. Costs of these entitlement programs depend on a combination of factors including program enrollment rates, benefit levels, and food costs, which factor in inflation. No major changes are made in eligibility requirements, benefit levels, or exemptions under the food stamp program. It is safe to say that food assistance expenditures will continue a trend toward relatively greater importance in U.S. Department of Agriculture spending as farm benefits fall and as the economy continues in recession.

Food Assistance Programs

Title XVII, now cited as the "Mickey Leland Memorial Domestic Hunger Relief Act," reauthorizes the food stamp program and related provisions. The food stamp program is reauthorized for five years (through FY 1995), retaining the current structure. The program retains measures that guarantee that households that meet eligibility criteria based on size, income, and assets automatically qualify for benefits. Benefits are still based on the "Thrifty Food Cost" plan as established by the U.S. Department of

Agriculture. The conference agreement authorizes amounts necessary to fund the program. Various House provisions that would have expanded potential eligibility and/or benefits were deleted in conference. However, provisions easing the way for the homeless, students, poor families receiving back-to-school allowances, rural residents who need mail delivery of stamps, and projects designed to increase participation by non-English-speaking minorities, the elderly, homeless persons, low-income working families with children, and rural residents are authorized (Democratic Study Group, U.S. House of Representatives, 22 Oct. 1990). Some food stamp program rules are simplified; additional penalties for fraud and misuse of food coupons are established (U.S. Congress, House of Representatives, news release). Additionally, state agencies are authorized to implement on-line electronic benefit transfer (EBT) systems in which household benefits are issued from, and stored in, a central data bank and electronically accessed by household members at the point of sale.

Other domestic commodity donation programs, including the Emergency Food Assistance Program (renamed from the previous TE-FAP, the Temporary Emergency Food Assistance Program) and the Commodity Supplemental Food Program (CSFP), are also reauthorized.

Commodity Programs/Food Costs

The new legislation freezes target prices and effectively decouples grain and cotton farmers' planting decisions at the margin, moving this part of agriculture a little farther toward market orientation. If a freer market results in more efficient use of resources, consumers should benefit modestly. On the other hand, sugar, dairy, tobacco, and peanut programs remain substantially nonmarket-oriented, with the consumer continuing to bear costs associated with market protection.

Specifically, target prices (1990) are maintained for wheat, feed grains, upland cotton, and rice throughout 1991-95 crop years. Because of flexibility provisions in the legislation, farmers are expected to alter crop choices on their flexible base of 15% of acres toward crops with higher net returns. In 1991, midwestern and northern plains farmers may shift some spring wheat acres into soybeans or minor oilseeds. Some midwestern corn acres may also be planted in oilseeds. If expanded soybean and other oil-

seed supplies result in lower market prices for protein, the effect on food costs would likely be felt through livestock prices. These could eventually reflect lowered costs of protein. Additionally, some of the edible oils could be lower priced. At this time, however, few estimates are available to predict how adjustments might be made to the new policies, and expectations are that effects on food costs will be minor. Ultimately, farm price outcomes for specific commodities will depend partially on how the secretary of agriculture sets acreage reduction program limits, whether to relax ARP acreage to foster competitive export prices or to maintain previous ARP acreage (U.S. Congress, Congressional Budget Office).

In the aggregate, retail food prices include a farm commodity cost component of approximately 30%, although this varies among commodities. Factors outside the aegis of the farm bill such as costs of labor and transportation (influenced by oil costs) often outweigh changes in the commodity cost of food in influencing food prices.

Food Safety and Quality

Consumer and environmental interest groups played an active role in the development of the recent farm bill. Early in the deliberations, a coalition of consumer and environmental groups released their *Farm Bill 1990: Agenda for the Environment and Consumers*. The agenda highlighted agriculture's effects on environmental quality including pollution of surface and groundwater, loss of wetlands, and deterioration and loss of natural resources. The report marked "excessive use of pesticides and nutrients" and pesticide residues in food as major concerns. Many of the concerns and potential solutions mentioned in this influential report in fact became discussion points and, in some cases, resulted in new measures or guidelines in the 1990 farm bill.

Organic Certification

For the first time, national standards for agricultural products labeled organic are established, which will give consumers who wish to choose products produced organically the opportunity. Standards apply to production and processing methods as well as to permitted materials, product handling, and labeling. Organic

certification has not been defined as a food safety measure per se, but rather a marketing measure that would organize the existing chaotic conditions applying to "organic" products. A determined coalition of producer and consumer interests succeeded in winning the measure against the long-standing adamant opposition of the U.S. Department of Agriculture.

Pesticide Recordkeeping

This new measure requires certified pesticide applicators to maintain records for two years on the use of restricted pesticides. These records are confidential but can be made available for the use of federal and state agencies and health care personnel.

Cosmetic Quality and Pesticide Use

The 1990 farm bill includes several new provisions that reflect society's ongoing concerns about the safety of pesticide and agricultural chemical use in agriculture. The issue of grades and standards, particularly those that pertain to surface—or "cosmetic"—appearance, was raised by several prominent consumer and environmental organizations during the 1990 farm bill debate. Their efforts resulted in a mandate to undertake several research projects to investigate the extent to which grades and standards, both voluntary and in conjunction with marketing orders, provide an incentive toward excessive pesticide use. The hypothesis is that cosmetic standards impede those producers attempting to reduce pesticide (and fertilizer) use and adopt alternative farming practices. Also mandated under the legislation were three two-year studies to evaluate marketing and education programs to provide consumers with choices of, and information about, commodities produced with alternative production processes. Whether these studies will be carried out by the U.S. Department of Agriculture and what they might lead to remains to be seen. Senator Dole amended the farm bill so that the numerous studies included in the 1,000-plus pages of the bill could be reduced to twelve and, reportedly, the cosmetic standards study currently is not on the short list.

In addition, the legislation permits citizens to petition the USDA for changes in grade standards. Previously, only the food marketing industry had standing for these purposes.

Country-of-Origin Labeling

Title XIII (fruits, vegetables, and marketing) contains a requirement to operate a two-year pilot program labeling products by country of origin.

Research

Research in general is treated fairly generously under the farm, appropriations, and budgeting bills. Agricultural research and extension programs are authorized for five years. The bills gradually increase funding authorization for USDA competitive research grants from \$70 million to \$500 million annually. The farm bill authorizes \$40 million annually for low-input research activities, \$20 million for integrated resource management research, and \$20 million for educational activities to extend practices to agricultural producers. Again, the differences between authorization and appropriation can be vast. Some suggest that an annual increase in competitive funding of about \$50 million is likely.

The GATT and Consumers

One goal of the GATT negotiations on agriculture was the "harmonization" of sanitary and phytosanitary standards internationally. Although it is extremely unclear at the time of writing whether any positive accord will result from the GATT, it is worth noting that some progress was made toward establishing a better framework for addressing trade disputes emanating from unlike national food safety and animal and plant health standards. Some agreement was reached to consider the scientific basis for food safety standards, and in so doing to provide a greater role to the three principle international food standards organizations. At this point, few countries are likely to agree to an international scientific court passing binding judgment in trade disputes, however.

Conclusions

Certain provisions of the farm bill respond to consumer and public concerns related to food safety and quality and the environment. Provisions that set stricter environmental compliance measures and penalties than previously (discussed in Antle's paper), reinforce the "green" theme of the farm bill. All in all, the bill moves agriculture and the public policy agenda farther along a road toward greater environmental responsibility and an examination of some of the unintended consequences of farm and marketing programs, grades and standards, and traditional research priorities. However, in the main, the farm bill retains much that is business as usual.

Several significant measures for consumers are not included in the 1990 farm bill. Some were passed separately, such as an extremely important nutrition labeling bill (P.L. 101-535 passed 8 Nov. 1990) and a nutrition monitoring bill (P.L. 101-455). Others remain unfinished business for the next Congress, including seafood inspection, FIFRA reauthorization, and the possible reintroduction of "circle of poison" legislation (see Natural Resources Defense Council).

References

- Environmental Scorecard for the 1990 Farm Bill: An Analysis of the Food, Agriculture, Conservation and Trade Act, with Recommendations for Future Action.* Washington DC: Natural Resources Defense Council, 1990.
- Farm Bill 1990: Agenda for the Environment and Consumers.* Washington DC: Island Press for the Center for Resource Economics, 1990.
- U.S. Congress, Congressional Budget Office. *Farm Program Flexibility: An Analysis of the Triple Base Option.* Washington DC, 1989.
- U.S. Congress, House of Representatives, Committee on Agriculture. "House-Senate Conferees Finalize Action on Farm Bill." News Release (PR90-91), 17 Oct. 1990.
- U.S. Department of Agriculture, Economic Research Service. "1990 Farm Bill Passed." *Agricultural Outlook*. Dec. 1990, pp. 32-45.

Farm Policy Reform and the Environment

John M. Antle

U.S. farm policy in the 1980s was marked by three significant developments: the international farm policy reform movement, the active participation by environment interests in the writing of the 1985 and 1990 farm bills, and the federal budget constraint. These developments raise new questions for farm policy. What are the environmental implications of domestic and international farm policy reform? Can both agricultural and environmental policy objectives be met in one piece of legislation? How does the federal budget constraint affect the feasible set of policy options? The goal of this paper is to outline an analytical framework that integrates technological, economic, and political factors for analysis of agricultural-environmental policy and to draw some preliminary conclusions about the environmental implications of the 1990 farm bill and trade liberalization proposals in the GATT negotiations.

Goals, Tools, and Tradeoffs

The principal goal of farm policy is to increase the welfare of farm interests using the available menu of policy tools, including production quotas, price supports, and acreage limitations. Producer welfare $PS(y)$ is defined as a function of output y and is measured as producer surplus. Absent market distortions, agricultural policies generate tradeoffs between producer welfare and the welfare of other groups in the economy, including consumer, taxpayer, and environmental interests. Let consumer/taxpayer welfare $CST(y)$ also be defined as a function of output y . CST can be measured as consumer surplus plus the total economic cost of taxation including dead-weight losses. For a given set of market demand and supply functions, $CST + PS$ is maximized at the market equilibrium output y_e . At all output rates greater than or less than the market

equilibrium output y_e , $CST(y) + PS(y) < CST(y_e) + PS(y_e)$, as shown in quadrant 2 of figure 1.

The goal of environmental policy is to increase the welfare of environmental interests, measured as willingness to pay for environmental quality. This component of social welfare is referred to here as environmental surplus and is assumed to be an increasing concave function $ES(EQ)$, where EQ measures environmental quality, as shown in quadrant 4 of figure 1. EQ is assumed to be a decreasing, convex function of farm output y , as in quadrant 3 of figure 1. EQ embeds the physical and technological relationships that translate production activity into changes in environmental quality through erosion, water contamination, and so forth. The EQ function shifts according to the efficiency of pollution prevention technology. This is a stylized representation of the relationship between production and EQ ; a more realistic representation would account for the heterogeneity of the physical environment at a disaggregate level of analysis (Antle and Just).

Let the curve in quadrant 2 represent the hypothetical case in which output is set costlessly by government fiat. The result is the efficient set OEC plotted in quadrant 1 of figure 1. Point E in quadrant 1 of figure 1 represents the market equilibrium where $CST + PS$ is maximized; ES is maximized at point C where output is zero. Actual policies that are not fully efficient generate points inside OEC . The area BEC contains all points corresponding to policies that reduce output below y_e . The area OEB contains all points corresponding to policies that increase output above y_e .

Policy Efficiency and Policy Reform

Policy reform can be defined as a movement toward more efficient policies. There has been much attention devoted to the question of efficient policy choice (Becker, Gardner, Tullock). This is a central issue in policy reform because inefficiency is the *raison d'être* of policy reform.

John M. Antle is an associate professor, Department of Agricultural Economics and Economics, Montana State University, and a University Fellow at Resources for the Future.

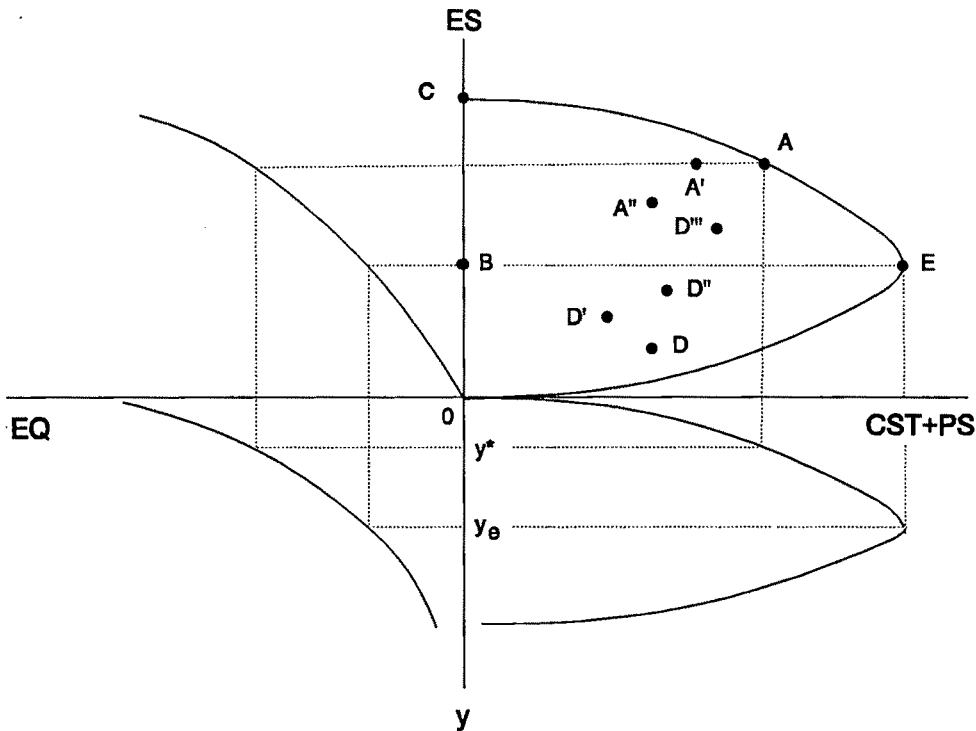


Figure 1. Welfare tradeoffs of agricultural policies

The socially efficient output level that maximizes aggregate welfare $CST(y) + PS(y) + ES(EQ(y))$ must satisfy $(CST' + PS')/ES'EQ' = -1$; that is, it will correspond to the point A on the efficiency frontier OE and an output level $y^* < y_e$. If income transfers to farmers are desired, lump-sum subsidies are the most efficient policy, assuming that the deadweight cost of taxation is less than that from other forms of intervention. If such subsidies were accomplished while holding output at y^* , the result would be a point such as A' in quadrant 1, where the difference $(A - A')$ measures the deadweight cost of taxation from the lump-sum transfer. The use of a production subsidy that increased output would reduce ES and entail higher deadweight losses and thus result in an even less efficient outcome such as point A'' .

A point like A'' is Pareto inferior to a point like A' , and it is tempting to conclude, as Becker does, that the political process should gravitate toward policies that are relatively efficient. Yet, relatively efficient subsidies, such as "decoupled" payments, are vigorously opposed by farm interests apparently because of their transparency. There is considerable evidence to show that many inefficient policies are used and persist over long periods of time in the United States

and elsewhere (Anderson and Hayami; Sturgis, Field, and Young).

There is, however, disagreement in the literature over the efficiency issue. This is a significant question for policy analysis because the prevailing paradigm of agricultural policy analysis is based on Pareto-efficient policy rankings derived from a welfare function or policy preference function (e.g., Rausser and Foster). The parameters of these functions are specified to reflect the political weights attached to competing interest groups. There are several reasons why policy rankings generated in this way may be erroneous. A potential Pareto improvement often is not politically feasible because the potential Pareto comparison of welfare gain does not account for political gains and losses to the politicians who make the policy choice (Antle and Johnson). The political process may prefer inefficient policies that obscure their redistributive effects (Tullock). And transactions costs in the political market may forestall efficient political trades.

The 1985 and 1990 Farm Bills

The conservation provisions of the 1985 farm bill—the CRP, sodbuster and swampbuster, and

conservation compliance—illustrate the growing influence of environmental interests in agricultural policy. In relation to figure 1, the 1981 farm bill was designed to support prices and expand output, thus putting U.S. agriculture at a point such as *D* in quadrant 1 of figure 1. The environmental lobbies were not able to eliminate all output-increasing aspects of farm programs, although the 1985 act did decouple target prices from variable production decisions by basing deficiency payments on program yields rather than actual yields. The 1985 act also had supply control provisions including the acreage reduction programs and the conservation programs. The CRP also transferred income to farmers, making it politically acceptable to both farm and environmental interests. Starting from point *D* in figure 1, this combination of policies moved the economy in a northwest direction to a point such as *D'*, raising environmental quality to some degree at the expense of consumers and taxpayers. While the 1985 farm bill did yield some environmental benefits, it is highly inefficient because it contains both output-increasing and output-reducing programs that work at cross purposes.

As the 1985 act was implemented its flaws began to become apparent. From an environmental point of view, the CRP was not efficient at improving environmental quality because its design did not target the most damaging erosion. As environmental interests focused more on potential chemical contamination of food and the environment, the inconsistencies between the output-increasing aspects of farm programs and environmental objectives were recognized. A related problem with the 1985 act, from the environmental point of view, was that the farm program participation rules limited a farmer's ability to rotate crops, thus possibly increasing the use of fertilizers and pesticides (National Research Council).

Another major problem with the 1985 act was that while it decoupled the target price from variable production decisions, it also transferred financial risk from farmers to the federal government by maintaining high target prices while reducing loan rates. The result was a program cost of some \$26 billion in 1986. The costs declined to about half that amount in 1988 and 1989 because of the higher market prices caused by the droughts in those years but remained high by historical standards.

Environmental interests were expected to be more influential in the design of the 1990 farm bill. A variety of proposals were put forward

early in the year, including better targeting and expansion of conservation programs (Center for Resource Economics et al.). Several factors diminished the environmentalists' influence, however, including disagreements among themselves over an agenda and the political strength of the farm lobbies in a year when all members of the House were up for reelection.

Early in 1990 the administration proposed to improve the economic and environmental efficiency of the commodity programs and reduce their budget cost by allowing farmers to plant nonprogram crops on their program base. Farmers would retain those acres in their program crop base but would not receive subsidies on those acres. This planting flexibility or triple base proposal was initially rejected by many members of Congress and farm interests apparently because, in their opinion, it transformed deficiency payments into politically unacceptable decoupled "welfare" payments. Planting flexibility was also opposed by some commodity producers who expected to be affected adversely as program participants diversified their production.

Environmentalists also sought a variety of regulations on land use and management practices, including tighter pesticide regulations and chemical recordkeeping requirements. These proposals were generally opposed by farm lobbies. Also contentious in the 1990 debate were the issues of wetlands conservation introduced in the 1985 act. Environmentalists felt the provisions were not being enforced and wanted them strengthened. Although this position was strongly opposed by farm interests, the environmentalists' position was strengthened by the president's campaign promise of "no net loss" of wetlands.

Without political pressure to reduce the budget costs of the farm programs, it is likely that the commodity and conservation provisions of the 1990 farm bill would have been very similar to those in the 1985 act, with somewhat expanded and better targeted conservation programs. Although the budget constraint imposed by the Gramm-Rudman-Hollings law imposed discipline on the options being considered during the first half of 1990, few anticipated the autumn budget crisis and the agreement between the administration and Congress to reduce farm program costs by an estimated \$13–\$15 billion over five years. This outcome favored the administration's farm bill proposal, as it precluded substantial expansion of the CRP or other costly conservation programs, and made the

administration's planting flexibility proposal politically feasible. The major budget-cutting provision of the bill thus became a mandatory 15% triple base. This outcome represents an increase in the economic and environmental efficiency of the farm programs. In terms of figure 1, U.S. agriculture will move from a point like D' to one like D'' as this policy is implemented.

The GATT and the Environment

At the time of this writing, it appears unlikely that a substantial measure of trade liberalization will be part of any final Uruguay Round agreement. This is unfortunate, as a reversal of the past ten years' trend toward agricultural protectionism would have been a significant accomplishment. A GATT agreement for agricultural trade liberalization would offer both problems and opportunities from an environmental point of view.

The key issue for environmental interests is whether trade liberalization would cause the United States to expand production in ways that will be environmentally damaging. Reduced export subsidies should cause world commodity prices to rise, encouraging production. But under a more market-oriented regime, only relatively productive land should be cropped. Environmental efficiency could be increased even as output increased if economically marginal lands are also the most environmentally sensitive. Thus, trade liberalization through a GATT agreement could bring about increases in both economic and environmental efficiency, corresponding to northwesterly movement from point D'' to a point like D''' in figure 1.

Economic and environmental efficiency also could be increased through the opportunity to redesign domestic subsidy programs (OECD). The GATT objective to decouple domestic subsidies from production could be made politically palatable to some farmers by converting price supports and deficiency payments into subsidies decoupled from production but linked to environmental objectives such as soil conservation, preservation of wildlife habitat, and reduced chemical use. This change in program design would face political difficulties, however. It would distribute program benefits differently than the existing programs, would not be attractive to agribusiness interests who favor output-expanding policies, and would be susceptible to budget pressures.

Trade liberalization makes some other aspects of environmental policy more difficult. Environmental regulations, such as restrictions on fertilizer and pesticide use, are likely to reduce the international competitiveness of U.S. farmers. Another difficult issue is harmonization of sanitary and phytosanitary regulations. U.S. environment interests oppose an agreement that would allow weaker food safety standards than current U.S. law.

Conclusions

Theories of government behavior suggest that policy reform will come incrementally unless there are substantial changes in political institutions and the influence of interest groups. They also suggest that when economic and political circumstances do change, there are opportunities for policy reform—policy changes that accomplish policy goals more efficiently.

Looked at from this perspective, the 1990 farm bill and the GATT should not be expected to be a "big deal." The degree of liberalization sought by the United States and the Cairns Group was not to be expected, as there is no dominant political constituency in the United States or among the GATT countries demanding policy reform to that degree. But neither do the 1990 farm bill or the GATT negotiations represent "business as usual." Although some of the most inefficient and inequitable aspects of domestic and trade policies still remain to be reformed, in some significant respects both the 1990 farm bill and the GATT present opportunities for increased economic and environmental efficiency. As such, both the 1990 farm bill and a possible Uruguay Round agreement can be regarded as positive developments in the evolution of agricultural policy design.

References

- Anderson, K., and Y. Hayami. *The Political Economy of Agricultural Protection: East Asia in International Perspective*. Sydney, Australia: Allen and Unwin, 1986.
- Antle, J. M., and R. N. Johnson. "Efficient Redistribution Reconsidered." Dep. Agr. Econ. and Econ. Montana State University, 1990.
- Antle, J. M., and R. E. Just. "Effects of Commodity Program Structure on Resource Use and the Environment." *Commodity and Resource Policy in Agricultural Systems*, ed. N. Bockstael and R. Just. New York: Springer-Verlag, 1991.

- Becker, G. S. "A Theory of Competition Among Pressure Groups for Political Influence." *Quart. J. Econ.* 98(1983):371-400.
- Center for Resource Economics et al. *Farm Bill 1990: Agenda for the Environment and Consumers*. Washington DC: Island Press, 1990.
- Gardner, B. L. "Efficient Redistribution through Commodity Markets." *Amer. J. Agr. Econ.* 65(1983):225-34.
- National Research Council. *Alternative Agriculture*. Washington DC: National Academy Press, 1989.
- Organization for Economic Cooperation and Development. *Agricultural and Environmental Policies: Opportunities for Integration*. Paris: OECD Publications, 1989.
- Rausser, G. C., and W. E. Foster. "Political Preference Functions and Public Policy Reform." *Amer. J. Agr. Econ.* 72(1990):641-52.
- Sturgis, R., H. Field, and L. Young. *1990 and U.S. Sugar Policy Reform*. Canberra: Australian Bureau of Agricultural and Resource Economics Disc. Pap. 90-4, 1990.
- Tullock, G. *The Economics of Special Privilege and Rent Seeking*. Boston: Kluwer Publishing Co., 1989.
-

Special Interests and the 1990 Farm Bill

B. Delworth Gardner

Antle and Kramer have written useful papers. Antle's contains an innovative analytical framework that is helpful in interpreting the changes in farm legislation discussed in both papers.

Antle's social welfare function that includes taxpayers, consumers, agricultural producers, and environmental interests takes us a long way toward understanding the package of agricultural policies that have emerged in recent years. Other interests exist, to be sure, such as foreign producers and consumers, and they too are active in our political arena. The crucial question, however, is who really counts in determining the final outcome of a given bill? My interpretation of the evidence is that domestic producers have the heavy political weapons, the metaphor of nuclear bombs comes to mind. The environmental interests have a tank or two and taxpayers and consumers perhaps a Molotov cocktail. Time after time during congressional debate on the 1990 FACTA bill, amendments that would have advanced consumer and taxpayer interests were beaten back by committee members who were loyal supporters of producers. Thus, Antle's statement that "the principal goal of farm policy is to increase the welfare of farm interests" is more valid than is implied by his model or explicitly stated in his discussion.

Kramer's long list of consumer interests promoted in FACTA cannot convincingly be used as evidence for dominating consumer power since almost all of them do not constitute a threat to producer wealth and income positions. Little political conflict is likely when a change in policy enhances the welfare positions of both parties. It is when some gain occurs at the expense of others that we discover who has real political clout.

Environmental interests did achieve something of a breakthrough with the CRP, sodbuster, swampbuster, and conservation-compliance provisions in the 1985 FSA bill. However, given inelastic demand curves facing producers,

the CRP can hardly be described as inimical to producer incomes, especially considering the direct transfer to producers of about \$55 dollars per acre placed in the reserve and the 50% cost-share provided by the government to establish plant cover. On 30 October 1990, Secretary Yeutter announced that fiscal year 1991 payments to about 330,000 producers participating in CRP would be \$1.6 billion dollars, or an average of \$4,848 per participant. This amount is not an insignificant reward for their participation.

Sodbuster and swampbuster provisions have clearly been onerous as well as inequitable to producers, given that all program benefits could be withheld from violators, and luck largely determined whether or not a particular farmer happened to be located on environmentally sensitive land. Consistent with my thesis of producer power, the Congress responded to these concerns in the FACTA by providing numerous exemptions to the penalties mandated in the FSA. I am not arguing that these exemptions necessarily injure consumers and taxpayers or that they even represent inefficient policy. But they are hardly what environmentalists and conservationists desired.

Antle points out that deficiencies in the environmental provisions of the 1985 act soon became apparent, and that better targeting of the real environmental problems is the remedy. Agreed. However, Congress has never been willing to tightly target agricultural programs, whether they be deficiency payments, credit programs, disaster payments, or conservation programs. Even the food distribution benefits discussed by Kramer have often gone to "unintended" beneficiaries. The reasons are obvious: the goal of every member of Congress is to maximize the benefits conferred on those to whom he/she is politically beholden. The result is a broad distribution of program payments and benefits across the demographic and geographic spectra.

The FACTA appears to mandate bold new initiatives in protecting wetlands and water

B. Delworth is a professor of economics, Brigham Young University.

quality. The voluntary water-quality program hopes to enroll 10 million acres of farmland through 1995. But, of course, landowners will be paid to implement water-quality plans. And, of course, cost-share payments for wildlife and wetland preservation are also authorized. In fact, the FACTA suggests that, if necessary, the government might pick up the entire cost of the wetland and land-easement initiatives. So, it is business as usual after all. Just as was the case with CRP, these new programs are sure to be utilized for income transfers to participants that are not justified by environmental gains and economic-efficiency benefits.

FACTA mandates that CRP acres may be targeted for water-quality and other environmentally sensitive problems in addition to soil erodibility. Further, highly erodible land criteria have been tightened and marginal pasture land can now be eligible for the CRP under specified conditions. While all of these environmental features of the new act may produce some benefits, they also will give more political constituents a claim on the budgetary resources of the government. I predict that an even greater proportion of the government's budget will be politically allocated rather than targeted to the most critical environmental problems. Congressional pressures will be mounted for a watering down if not an evasion of those provisions that cut into producer profits.

Antle points out that Becker has argued that income transfers to producers will be done in ways that minimize deadweight efficiency losses.

Decoupling of payments from production decisions has many efficiency and environmental advantages and thus would reduce deadweight losses. However, less than 10 million acres of land have thus far been idled in the 0/92 and 50/92 decoupled programs as compared to about 30 million in the CRP and around 50 million in annual acreage set-asides. So much for efficient transfers! But why? Because producers have the political clout to get their transfers built into support prices, acreage allotments, and internal and external quotas where the payments are largely hidden from consumers and taxpayers.

FACTA has extended the marketing loan to oilseeds and possibly to wheat and the feed grains to guarantee that American exports are competitive. The budgetary costs with the taxpayers footing the bill could be enormous if the EC continues to subsidize exports, keeping world market prices low. The implications for the environment are unclear depending on the quantities and locations of acreage that is cropped and the incentives to use chemicals.

The triple base option gives farmers more planting flexibility in maintaining program crop bases than was available previously. As Kramer points out, this will probably result in more oilseed production and lower consumer prices for cooking oil and perhaps slightly lower livestock prices. But this could be offset by higher market prices for the grains which are displaced. Kramer seems to have it about right when predicting only a slight impact on consumers. Too bad we cannot say the same for taxpayers.

The 1990 Farm Bill and the Uruguay Round: Discussion

Daniel A. Sumner

The papers by Runge, and Rossmiller and Nugent review some major U.S. commodity policy changes of the last year (the 1990 Food, Agriculture, Conservation and Trade Act and the 1990 Budget Reconciliation Act). They then consider the major prospective policy change for international agriculture, the still-evolving Uruguay Round of multilateral trade negotiations under the GATT. This discussion follows the same path, and the main objective is to clarify some points raised in the two papers.

Most of the Runge paper consists of speculation on the politics of the process of creating the 1990 farm legislation or on the internal motives of the individual actors. But every participant or observer will have a different set of impressions. For example, from my perspective it seems clear that much of the core of the Administration Green Book was included in the final act. Instead of continuing in this vein, let us turn to the results for commodity policy.

For target price commodities (grains and cotton), payment yields were kept frozen for the next five years and effective loan rates were raised slightly but kept well below the range likely to provide a floor to market prices. Keeping these features of the 1985 act avoided a return to the major distortionary aspects of earlier programs. The innovation of the 1990 act was that the budget reduction (of some \$11.2 billion in projected commodity program spending over five years) was achieved in ways that should enhance the economic performance of the sector. Outlays were reduced by reducing the base acres eligible for deficiency payments by 15%. Under the new programs, even for program participants, the anticipated market price is the relevant supply price for marginal acreage decisions among program crops and oilseeds. With frozen program yields implying that market price is the relevant supply price for producers on the intensive mar-

gin as well, target price programs are substantially less distortionary than just five years ago. Farmers face market prices for all marginal production decisions. Further, the provision of the 1990 legislation that protects an additional 10% of base if planted to alternative crops allows more short-run supply response to market conditions by eliminating the potential capital loss of valuable base for changing crops. Unlike the lowering of target prices in the 1985 act, the structure of the programs themselves was made more market oriented in 1990.

One of the major controversies in 1990 was over eliminating program payment eligibility for those associated with farms with large gross sales or for those participants with high net incomes. These and other targeting schemes were pursued most vigorously outside the agricultural committees. Targeting raises basic questions about the purpose of commodity programs. If they are really income maintenance programs for the deserving needy, then the distribution of payments is indeed odd (Sumner). Alternatively, if their role is to regulate markets as industrial policy, then perhaps direct payments to individuals is not the appropriate program structure. In the end, no targeting was enacted in 1990 and little progress was made in clarifying the issues.

Commodity programs without direct payments did not progress toward market orientation in 1990. In 1990 the legal details of U.S. sugar import program was changed in response to a GATT challenge from Australia, but this was done administratively and will likely have no significant effect on the economic operation of the program. Legislation left the sugar (and peanut) program mostly unchanged from 1985.

Dairy policy saw a clear move away from market orientation in 1990. The 1985 Food Security Act provided a formula that keyed support price changes to the level of government purchases of surplus production. The support rate came down after 1987 and was likely to continue a slow downward path with continued expected surplus for the next year or two. But the 1990 act severed the link to market forces and

Daniel A. Sumner is the Deputy Assistant Secretary for Economics, U.S. Department of Agriculture, and a professor, Department of Agricultural and Resource Economics, North Carolina State University.

froze the support price at the current \$10.10 per hundredweight. The result will almost surely be large surpluses and outlays. Provisions for producer assessments and some unspecified suggestions for supply controls are what the legislation offers to stem the anticipated milk glut.

One provision of the dairy assessment scheme should be highlighted for its perverse implications. For the initial 5¢ per hundredweight assessment (scheduled to increase), dairy producers who do not increase their output from the previous year will have their assessments refunded. To qualify for a refund requires no growth, and being just one pound over the no-growth level requires payment of the full assessment. Consider the producer who sold 15,000 hundredweight at \$12 in 1991 and faces a 20¢ assessment for 1992. In this case, if she sells 15,001 hundredweight at \$12 per hundredweight in 1992, her net of assessment receipts will be \$2,988 less because of that last hundredweight of milk. For a 10% expansion the marginal net price for the last 150 hundredweight is not \$12 or \$11.80 but \$9.80 and increases asymptotically toward \$12. The demand price facing each producer has a discontinuity that implies an effective choice between a large expansion or no expansion at all. Young producers are penalized as are those with improving efficiency or other incentives to expand. The winners are those who are already well established and secure, though for all producers the incentives for herd improvements and other efficiencies are reduced.

The Uruguay Round

The Uruguay Round of trade negotiations was still underway in late December 1990, and the

months between the writing and printing of this discussion should resolve many of the uncertainties about the outcome of the Round. In this context perhaps two points are useful.

First, in the realm of reasonable outcomes for agricultural trade, our analysis at the U.S. Department of Agriculture (USDA) indicates no additional internal support reductions (above what was accomplished in 1985 and 1990) for the target price commodities in the United States (except perhaps for rice). The United States has made major cuts in internal production-related support levels since 1986, and the payment base reductions and other changes in 1990 are larger than other countries have been able to match. By suggesting further reductions in deficiency payments are necessary in a GATT agreement, Runge has simply not analyzed the proposals carefully.

Second, the set of outcomes based on the Hellstrom proposal in Brussels includes options that can have clear, measurable benefits for U.S. agriculture and trade. Rossmiller and Nugent are overly pessimistic that trade gains are not in the set of a potential compromise. Even a 30% reduction in European Community subsidized exports over five years would raise grain and other commodity prices substantially and provide new market opportunities for the competitive exporters. This is clearly on the path to bringing discipline to agricultural trade and should be enough to garner the support of commodity interests and economists alike.

Reference

- Sumner, Daniel A. "Targeting and the Distribution of Program Benefits." *Agricultural Policies in New Decade*, ed. Kristen Allen, pp. 125-52. Washington DC: U.S. Department of Agriculture annual policy review, 1990.

Data Needs to Assess Environmental Quality Issues Related to Agriculture and Rural Areas

Jerald J. Fletcher and Tim T. Phipps

The assessment of environmental quality problems and the design and economic evaluation of alternative federal, state, and local policy strategies to address environmental issues require complex analyses which must address the underlying physical, biological, social, and economic processes. The study of environmental problems is thus inherently multidisciplinary and should involve the physical, biological, and social sciences. Unfortunately, the multidisciplinary nature of environmental issues has caused problems with the quality and general availability of the data necessary for appropriate analyses. While data have been collected in many areas to address discipline-specific problems, there has been little regard for the eventual need to coordinate analyses across disciplinary lines. As Portney points out, this lack of coordination in data collection is reflected in the often Balkanized system of data collection among federal statistical, research, and action agencies.

The activities of the Economic Statistics Committee (ESC) of the American Agricultural Economics Association (AAEA) highlight the need for the appropriate data to analyze both the causes and consequences of environmental problems as well as the set of alternative policy solutions that have been or may be proposed to alleviate such problems (Bishop and Ervin).¹ This paper reflects the continuing concern of the ESC with these issues.

Jerald J. Fletcher and Tim T. Phipps are associate professors in the Division of Resource Management, West Virginia University.

This is Scientific Article No. 2281 of the West Virginia University Agriculture and Forestry Experiment Station.

Helpful comments by Virgil Norton on an earlier draft of this paper are acknowledged.

The ESC sponsored a survey during 1988–89 followed by a workshop, "New Directions in Data, Information Systems, and Their Uses," prior to the 1989 AAEA annual meetings in Baton Rouge, LA, to develop a comprehensive picture of the current and future data needs of agricultural and rural social scientists, (Hushak, Chern, and Tweeten).

This paper builds on previous discussions of the methodological approaches to, and the associated data needed for, the study of environmental problems. The primary focus is on the underlying issues that must be resolved in this area, not an exhaustive review of current data availability nor a repeat of the common call for additional data and data collection efforts. The promise of relatively recent advances in computer methods (primarily geographic information systems—GIS) for data development and analysis and the implications for future directions for empirical economic analysis of environmental problems are discussed. There is also discussion of a variety of separate, but related, issues that pertain to environmental data needs. Finally, we derive implications from the discussion that can guide current and future data-gathering efforts to aid in the analysis of environmental problems and policies.

A Framework for Assessing Agriculturally Related Environmental Data Needs

To gain a better perspective on the data needs problem, it is useful to consider the framework in which the data are to be analyzed. Just and Antle develop a conceptual framework that considers interactions between agricultural and environmental policies. The framework is based on a model that shows that farmers' production decisions generate a joint distribution of output, input, environmental attributes, and pollution. Using their framework, they argue for the need to know the joint spatial distribution of both physical and economic factors in order to analyze the economic and environmental consequences of agricultural and environmental policies. They conclude: "Our analysis points to the kinds of data that are needed to make valid in-

ferences. Statistically reliable field-specific production data and environmental data would make possible measurement of key parameters (such as the correlation between production decisions and environmental attributes of land) that are needed to assess the aggregate relationships between agricultural policy, environmental policy, and the environment" (p. 202). We second Just and Antle's call for greater coordination of the research and data-gathering efforts of the various disciplines involved in the analysis of environmental problems and policies. The need for information on the joint distribution of economic, social, physical, and biological parameters should be recognized. We would, however, like to discuss the implications of their conclusions for the further development of appropriate data and analytical methods.

In practice, the collection and analysis of data relevant to environmental policy problems does not, and should not necessarily, proceed only from the field level to the aggregate. The development of the appropriate data is a dynamic process that iterates between aggregated analysis of the data at hand and more detailed data collection based upon the results of prior aggregated analysis. Analysis of the additional data and the implications gleaned from further aggregate analysis may lead to another round of data gathering as well as, in many cases, redefinition of the policy problem or issue under consideration. That is, data needs are fundamentally related to the problem being analyzed; the problem defines the data needs, while the availability of data limits the problems that can be addressed. The Just and Antle framework should not be interpreted to mean that the availability of field-specific data is a necessary requirement for spatially detailed analysis; the development of such data is neither necessary nor economically efficient for many policy problems. As their framework makes clear, it is the correlation of relevant factors that is important. The level of detail necessary in the data collection that this implies largely will depend on the spatial and temporal correlation coefficients among the factors.

Geographic Information Systems

Geographic information systems (GIS) have the potential to greatly increase the usefulness of existing data in the analysis of environmental problems. A geographic information system is a combination of computer hardware and soft-

ware designed to collect, manage, analyze, and display spatially referenced data. The GIS approach to handling vast amounts of data while preserving spatial detail has been enthusiastically received by many of the scientific disciplines and agencies working in areas germane to environmental issues. Examples of GIS applications include such diverse activities as describing and managing natural resources such as forests, soils, water, and minerals; use by utility companies to plan routing and service areas; application to the work of local area planning committee, tax assessors, zoning boards, surveyors; and the list could be extended. Because of the large number of observations and the complexity of the relationships among the data points necessary to provide spatially detailed information on resource characteristics that vary significantly over space and/or time, computer applications of GIS have traditionally been expensive and reserved for relatively high-powered mainframe systems with large amounts of data storage. Recent advances in the computing power of personal workstations has made computer-intensive data analysis such as GIS cost effective.

Economists, however, have not yet widely adopted GIS approaches nor adapted empirical economic analysis to utilize the potential benefits of GIS. While GIS and related technologies are making it possible to generate, manipulate, understand, and disseminate information about the world far more broadly and pervasively than has been possible with previous technologies, economists have often seen only the final results—the maps or other graphic displays of physical phenomena. Such a picture hides the individual features of the many thousands of data points of which it is composed. GIS may be incorrectly perceived as merely a useful tool for creating maps, not as a separate tool to be used in spatial analysis.

A discussion of the potential that GIS hold for strengthening economic analyses may open new research paths in economics. GIS can significantly improve our ability to analyze environmental problems using conventional approaches by significantly improving the quality and the quantity of available data. In addition, we believe that GIS will open new avenues of empirical analysis based on spatial statistics and result in the development of novel and innovative approaches to policy analysis. The first issue involves new tools for old questions, while the latter may be summarized as new questions arising from new tools.

Using GIS to Facilitate the Use of Available Data

Numerous government agencies devote a considerable portion of their budgets and available manpower to collecting information on the world in which we live. Environmental data (or data relevant to the analysis of environmental problems) are collected by diverse agencies in many departments. For example, in the Department of Commerce, the National Oceanic and Atmospheric Administration has extensive spatial and temporal data on weather patterns and climatic conditions, while the Bureau of the Census develops extensive information on many demographic and economic aspects of our industrial society. In the Department of Interior, the Geological Survey has data on ground and surface water quantity and quality, on geologic formations, chemical compositions, and many other measures of interest to specific problems; the Bureau of Land Management maintains information on portions of the public lands, and the Fish and Wildlife Service has data on wildlife and fishery habitat and abundance as well as recreational use. In the Department of Agriculture, the Soil Conservation Service has detailed information on soil characteristics and related data for private lands; the National Agricultural Statistics Service (NASS) develops information on agricultural inputs, production, and related items; and the Forest Service has information on recreation use, fisheries, and many other aspects of public forest lands besides those directly related to timber production. The Environmental Protection Agency has extensive data sets on pollution sources. The National Aeronautics and Space Administration and the Departments of Defense, Energy, and Health and Human Services all collect and maintain data that may be useful in analyzing specific environmental issues. This list reflects the diverse nature of the data useful in the analysis of environmental issues. It could be extended significantly, but the point is clear: the sources of relevant data are numerous and diverse. (See Yarden et al. and Garkey and Chern for specific details on data availability.)

The various collections of data for analyzing environmental issues are quite diverse, and published statistics are of such an aggregate nature that they often are not useful in developing insight into the spatial correlations and complex interactions between the physical resource base, natural phenomena, and man's activities that result in environmental problems. The raw data on which such statistics are based, however, often

are of sufficient detail and contain the temporal and spatial reference points necessary to develop linkages between disparate data sets. Such data can be used in the analysis of problems not contemplated by the original collectors of the basic data. In many instances, however, while the basic information may be included in the raw data, the form and method of storage varies so greatly across agencies and among data sets that systematic joint analysis of the disparate data sets has been difficult and costly at best and often virtually impossible from a practical perspective. Recent developments in computer hardware and software may provide a key to better data utilization if institutional and legal constraints can be overcome.

Much of the information available on the resource base has yet to be included in the economic analysis of environmental problems, although aggregate statistics drawn from the data are often used. In many cases, the information on resource characteristics has been reported as maps, often with the information summarized in a series of contour or "iso-value" lines. Examples include average rainfall, elevation, precipitation, soil type, temperature, habitat, and so on. The ranges reported are related to the scale of the map and the properties underlying the spatial end, in some cases, temporal distribution of the characteristic of interest. Using such information in even a simple economic analysis has traditionally required a large investment of both time and effort in the development of a data set that reflects the attributes of the resource characteristic and is of a scale or level of aggregation suitable to the problem under consideration.

In recent years many of the older maps have been stored in electronic format as part of a GIS system, and newer maps are often a visual depiction of information stored in digital form. If the information is available in a suitable GIS format, summary statistics relative to a single resource variable or combination of variables that includes spatial references can be developed with relative ease. GIS systems can also be used to combine synthetic data, for example estimates of nonpoint pollution loading from agricultural lands produced by simulation models (perhaps based on observed parameters of the resource base and hypothetical or observed management parameters contained in a GIS), with observations on point sources to obtain overall estimates of an environmental quality parameter such as BOD in ambient water. The data requirements for the sound analysis of environmental issues may be best met from nontraditional sources.

Thus, current GIS systems can increase the number of data series available for conventional economic analysis if the economist knows about the relevant data and has access to it. With the increasing prominence of multidisciplinary approaches to problems that cut across traditional disciplinary boundaries, it is important that existing data originally gathered to support research in individual disciplines be catalogued and made accessible to others.

Using GIS to Develop Innovative Approaches to Economic Policy Analysis

Of more interest is how GIS and related technical developments in statistics may provide new paths for empirical analyses. As a profession, economists have concentrated on time-series analysis, the analysis of data that have temporal reference points but without spatial dimension. When the concept of space has been introduced in the analysis, it has usually been in a simple way, such as the way distance is introduced in travel cost models to measure resource price. While developments in spatial econometrics have been reported in the literature (see Paelinck and Klaassen for an introduction), applications have been primarily in the regional science literature and heavily dominated by Western European citations. The promise that GIS hold for the improved analysis of environmental problems rests largely on the ability of GIS to provide the foundation necessary to develop spatial autocorrelation functions and other measures of spatial correspondence that may be important in gauging the effects of environmental problems and alternative policies.

Issues Related to Environmental Data Needs

A number of specific issues are especially pertinent to the data needs for evaluating environmental problems. These include barriers to the accessibility of existing data, such as confidentiality, moves toward the privatization of data, the effects of nonmarket aspects of critical factors related to environmental quality, uncertainty in the data, and the appropriate level of reliability.

Barriers to Access

A number of difficult issues are related to institutional and legal barriers to data access. One

of the most pressing is confidentiality—an individual's right to privacy versus society's right to know. Confidentiality is particularly a problem with environmental data because of the need to preserve spatial detail in the analysis and often in the presentation of results. Confidentiality is often dealt with in other areas by making the data available only in an aggregated form. Unfortunately, as discussed above, aggregated data are often not useful in analyzing environmental problems. This issue will continue to receive attention from a broad spectrum of interests, and the final outcome is far from decided.

A second, but related, issue is duplication of effort. While duplication has traditionally arisen from lack of communication among those gathering the data; now it sometimes arises because of confidentiality constraints. If agencies are unable to share information because of confidentiality rules, both may gather the same basic information to meet their stated objectives even if additional burden is placed on respondents. Again, a tradeoff occurs between the benefits of privacy and the costs of duplication.

A third issue is the move toward privatization of data. For example, much of the remote sensing data is now obtained from SPOT, a for-profit corporation whose objective is to make money from the images collected. Another and more insidious problem is the move toward private distribution of data collected by public agencies. For example, various private companies market value-added products based on census data. While this is in many cases appropriate, if private provision of data leads to reduced public access to data, this seems to undercut the primary reason for public data collection efforts, the public good aspects of information, and the natural monopoly nature of data collection efforts.

Security and national defense comprise the fourth issue. The information available via remote sensing, satellite photography, and other modern technologies is of vital importance to national security interests and therefore is often classified. If made available to the general public, on the other hand, the information obtained by security and defense agencies could dramatically alter the information available for the analysis of environmental and resource issues.

Effects of the Non-Market Characteristics of Environmental Issues

As noted by Bishop and Ervin, many environmental issues involve actions or goods that are not traded on markets and cannot be measured

by observing market transactions. Analysis of environmental issues must, therefore, often rely on nonmarket techniques to measure economic benefits and costs. In many such cases, economists are left to collect the necessary information directly. Many of these nonmarket techniques, such as contingent valuation, utilize survey methods that are not generally part of the training of most economists (Bishop and Ervin). This need for greater training in nonmarket evaluation techniques was driven home by the level of attendance at the nonmarket benefit session at the Vancouver meetings in 1990.

Perhaps more important from a data perspective, the data generated as part of a nonmarket benefit study are often relevant only to the particular situation under study. The development of methods that would circumvent this problem could return large benefits. One possibility is to develop and specify models to evaluate nonmarket inputs and outputs that allow generalization of the estimated parameters to other regions and situations (Walsh, McKean, and Johnston).

In any case, economists cannot look to our fellow researchers in other disciplines to solve the problem of valuing nonmarket goods. Valuing environmental amenities, damages, and other nonmarket goods belongs to the economist. The basic understanding of these nonmarket values is not yet sufficient to allow for broad collection of appropriate values by a statistical agency. However, it is important to use the information gleaned from specific studies to develop methodologies that will complement the information available on the resource base. The result will better guide policy development and evaluation.

Uncertainty in the Data

Bishop and Ervin also note a tendency for greater uncertainty associated with natural resource and environmental data than with most data used in other areas of research. Each step in what is often a complex chain of interactions that result in environmental problems has, at best, some uncertainty associated with expected outcomes. As these steps are combined, the uncertainty may be compounded. However, because coefficients and/or data developed by other disciplines are often entered into an economic analysis as nonstochastic terms, the true nature of the uncertainty in a particular problem may be obscured.

Uncertainty is associated with the location and behavior of a pollutant in the environment. For

example, it is often necessary to know how pollutants are transported spatially, potentially a three-dimensional problem. In addition, there is an added temporal dimension of decay or accumulation that depends on the environmental characteristics of each point in space (e.g., soil type) and time (e.g., temperature). Thus, uncertainty enters not only with location of a pollutant in four-dimensional space-time but also with the environmental attributes and chemical processes associated with each point. Locational and behavioral uncertainty is compounded by uncertainty of the effects. That is, not only is the level of pollution uncertain, but the environmental or health damages caused by the pollutant are often only partially understood. This uncertainty is compounded when the secondary effects from combinations of pollutants, intermediate compounds formed by the pollutant, or (once ingested) metabolites of the pollutant are considered.

The great complexity and uncertainty associated with the location, behavior, and damages of a pollutant mean that information is required from a wider variety of sources than is necessary for most problems analyzed by economists. Frequently, the available data were originally developed from samples fraught with problems of sample selection bias. That is, much of the available data is not based on random or stratified random samples but is, as noted by Bishop and Ervin, driven by particular problems, public complaints, or "hot spots." Typically, the available environmental data relevant to a particular problem will be a combination of randomly sampled data for certain regions, no data for other regions, and biased or "hot spot" data for still other regions. While data collected on a nonrandom basis contain less information than a well-constructed random sample, it does contain some information. The trick is to combine the available information so that the statistical integrity (or lack thereof) of each sample is preserved, and then to present the results in a way that accurately conveys the differing statistical reliability for each region.

Appropriate Level of Data Reliability

As mentioned earlier, the level of accuracy or reliability of the data needed is related to the problem under consideration. For water pollution from agricultural production, data at the field level or even lower may be necessary to predict accurately both the economic and environmental

effects of alternative policies if relevant field characteristics vary significantly. While this level of spatially detailed data would be ideal, it is not economically efficient or necessary to gather such data to analyze many issues of interest. While this level of detail would be needed if the policy were a centralized command and control program, most of the environmental problems related to agriculture are nonpoint in nature and not well suited to such policies. Policies that address agricultural nonpoint pollution are often handled more efficiently by a mix of centralized and decentralized decision making, with each level of decision requiring data of different reliability or scale of sampling.

Soil conservation policy provides a good example. As part of the National Resources Inventory (NRI) first conducted in 1977, data on soil erosion were developed that demonstrated that erosion problems are highly concentrated in a few regions of the country. This information spawned the policy problem of targeting soil conservation spending to those regions with the most severe soil erosion problems. While such targeting is more cost effective than not targeting, there is also high intraregion variability in soil erosion rates. Unfortunately, while the NRI is statistically valid at the major land resource area level, it is not detailed enough to determine the actual fields, or portions of fields, that require treatment.

The NRI, however, can be used in two ways to guide soil conservation policy. The first is to help policy analysts determine which regions of the country, and the characteristics of areas within each region, where it would be most efficient to gather the highly detailed data needed for field-level treatment. A second, and related, possible use is to serve as the first step in a multistage budgeting process, where the first stage is the centralized decision of how to allocate the soil conservation budget among competing regions and the second is the regionally decentralized decision of how to allocate regional budgets among farms. The third step is fully decentralized decisions made by individual farmers about how to allocate received soil conservation funds among competing techniques and fields. This step would necessarily require detailed knowledge of the resource base as well as characteristics of the farm firm, but attempting to use this information in the initial allocation stages may not be practical. It is still likely, however, that the information necessary for local implementation of any policy or plan will be much more detailed than that necessary to guide the formulation of national policies and guidelines.

Conclusions and Future Implications

We have not attempted to reiterate the lists of environmental data needs that have been put forward in numerous pleas for more data. While more data are needed—in some areas the information available is woefully lacking—there is a great need for consolidation and assessment of existing data to guide additional data collection efforts. While the need for additional data is obvious in many areas, in other cases the need could be filled, at least in part, by existing information housed in seemingly unlikely places. Many agencies have gathered considerable data on the resource base of the United States—we probably know more about our country's resources than any other in the world. However, much of this information is virtually inaccessible to those most interested even if they have managed, by some happenstance, to learn of it. We can only reiterate our perceived need for a clearing house for environmental data, a "Bureau of Environmental Statistics" as so ably stated by Portney.

Such a central agency for the collection, evaluation, and delivery of critical environmental information would provide a tremendous leap toward meeting environmental objectives. A central environmental clearing house could provide for combining data sets (perhaps individually confidential) from diverse sources in ways necessary for appropriate analysis while still maintaining the required levels of confidentiality and protecting the rights of the individual. Appropriately funded and supported, a central clearing house could coordinate data collection efforts, reduce duplication, and provide guidelines for a common data format.

Interdisciplinary communication on data collection and collaboration across disciplines on experimental design also are needed. While disciplinary studies may concentrate on particular aspects of a problem, a coordinated effort could develop not only discipline-specific data but add those data necessary for the results and conclusions to be assimilated in other research by other disciplines. For example, agronomic studies based on an experimental design that creates the data necessary for analysis of variance tests of alternative levels of a particular treatment may provide little insight into the tradeoffs and substitutions among inputs. Many such studies have not recorded the levels of factors held as controls within a particular experiment but necessary to make the results useful in understanding broader questions related to input controls and other factors of interest.

We would hope that current and future data-gathering efforts will result in an integrated approach to data collection and analysis that will allow for the compilation of the spatial and temporal correlation functions necessary to guide aggregation of site-specific studies to broader aggregates. Examples of such efforts are the crop-specific studies conducted by NASS in cooperation with the Economic Research Service (ERS), the Management Systems Evaluation Area (MSEA) projects that result from cooperative efforts among a variety of agencies concerned with environmental problems in agriculture, and the area studies anticipated by the ERS. While we cannot expect all data needs to be filled immediately, careful, coordinated planning of efforts now underway can significantly increase the effectiveness of such studies in shedding light on underlying data and aggregation issues.

References

- Bishop, Richard, and David Ervin. "New Directions in Data, Information Systems, and Their Uses: Natural Resources and the Environment." *Proceedings of the Symposium on New Directions in Data and Information for Rural Areas*. Ames: Iowa State University Press, forthcoming.
- Garkey, Janet, and Wen S. Chern. *Handbook of Agricultural Statistical Data*. Dep. Textiles and Consumer Econ., University of Maryland, 1986.
- Hushak, Leroy J., Wen S. Chern, and Luther Tweeten. "Priorities for Data on Agriculture and Rural Areas: A Survey of Agricultural and Rural Social Scientists." *Proceedings of the Symposium on New Directions in Data and Information for Rural Areas*. Ames: Iowa State University Press, forthcoming.
- Just, Richard E., and John M. Antle. "Interactions Between Agricultural and Environmental Policies: A Conceptual Framework." *Amer. Econ. Rev.* 80 (1990): 197-212.
- Paelinck, Jean H. P., and Leo H. Klaassen. *Spatial Econometrics*. Westmead, Farnborough, Hants., England: Saxon House, 1979.
- Portney, Paul R. "Needed: a Bureau of Environmental Statistics." *Resources* 90(1988):12-15.
- Walsh, Richard T., John R. McKean, and Don M. Johnston. "What Can We Learn from 20 Years of Work with TCM and CVM?" *Benefits and Costs in Natural Resources Planning*, ed. John P. Hoehn. W-133, Third Interim Rep., Michigan State University, 1990.
- Yardas, Davis, Alan J. Krupnick, Henry M. Peskin, and Winston Harrington. *Directory of Environmental Asset Data Bases and Valuation Studies*. Washington DC: Resources for the Future, 1982.

Data Needs to Address Economic Issues in Food Safety

Tanya Roberts and David Smallwood

Invisible hazards and imperfect information held by producers, consumers, and regulators continue to make food safety a volatile issue that can disrupt markets and cause substantial losses to suppliers of agricultural inputs, farmers, marketers, and consumers. New and improved data are needed to assess risks, consumer demand, and costs of alternative private and public control strategies.

Food safety data were identified as a high priority by several groups. The 1989 Survey of Agricultural and Rural Social Scientists ranked "very important" data collection and dissemination efforts and gave "develop measures of food safety" a rank of seventh and "inventory of pesticide and herbicide applications" a rank of ninth (Hushak, Chern, and Tweeten). A report of the planning and budget subcommittee of the Experiment Station Committee on Organization and Policy (ESCOP), *Research Agenda for the 1990s: A Strategic Plan for the State Agricultural Experiment Stations*, rated "ensuring the safety and stability of consumer foods" as its first priority. The National Research Initiative Competitive Grants Program in the U.S. Department of Agriculture (USDA) added food safety to its list of program areas for funding in 1990, largely because of ESCOP's rating. President Bush proposed a major food safety initiative in 1989 with a large data collection component, and Congress appropriated funds in 1990.

Better data are needed on the incidence, severity, and economic dimensions of foodborne disease from microbial (bacteria, parasites, fungi, and viruses), chemical (insecticides, herbicides, fertilizers, animal drugs, environmental con-

taminants, food additives), and natural constituents (including biotechnology) in food.¹ No data source exists with definitive estimates of the number of illnesses caused by foodborne sources or the distribution of disease severities (OMB). The estimates in the literature suggest that microbial sources are causing from 6.5 to 33 million cases of foodborne disease annually and 9,000 deaths (Bennett et al., Garthright, Archer, and Kvenberg). Chemical sources of foodborne disease risk are less well characterized. A worst case EPA estimate suggests 6,000 cases of cancer annually (EPA 1987), while others have reckoned the risk to be much smaller (EPA 1990, Ames and Gold, Archibald and Winter). Economists can contribute to the food safety discussion by identifying data needed for economic analysis, estimating the social welfare costs of foodborne illness, and estimating benefits and costs of alternative control programs.

More Information Improves Social Welfare

The central food safety problem is lack of information. Consumers cannot look at a package of food and see either chemical residues or microbial contaminants. Food safety attributes are difficult for consumers to quantify even after consumption or "experience" with the product. For chemical residues the human health effects, such as cancer, may not be experienced for many years. For many microbial risks, consumers cannot correlate flu-like symptoms with food because the lag of days to weeks between food consumption and illness prohibits identification of the specific food or other source of illness. Thus, food safety attributes are not readily apparent either before purchase or after consumption of food.

Producers of risky foods do not have an in-

Tanya Roberts and David Smallwood are agricultural economists in the Food Policy Research Section, Economic Research Service, U.S. Department of Agriculture.

The views expressed in this paper are solely those of the authors and should not be interpreted as official policy of the U.S. Department of Agriculture.

Review comments are appreciated from Betsy Frazao, Lester Myers, and John Schaub, Economic Research Service; Peter Kuch, Environmental Protection Agency; Bev Fleisher, National Science Foundation; David Berkowitz, Food Safety and Inspection Service; and Bing Garthright, Food and Drug Administration.

¹ There are two emerging areas of data needs that we have not been able to address: biotechnology and natural toxic constituents. Ames and Gold believe that the human health risk from natural plant constituents is several times greater than the risk from chemical residues in foods.

centive to disclose risk information because they will make higher profits if consumers are uninformed. As Akerlof demonstrated, asymmetrical information about product attributes creates market failure.

Market forces not only disclose too little food safety information but also generate too little information. Basic research about food safety, such as developing risk assessment mechanisms or new tests to identify foodborne hazards, is underfunded because of its public good characteristics.

One could imagine that, even if the market did not disclose information optimally, private liability laws could provide the necessary incentives for firms to produce the socially optimum level of product risk. However, Viscusi demonstrated that tort liability does not work efficiently to redress health and safety problems in the marketplace because only a subset of the total harms are compensatable, and only risks which have discrete, easily traceable causes are handled well by the tort system. An efficient product liability system can sometimes remedy the disclosure of information problem but cannot solve the general lack of information problem.

Government subsidization of basic research and data collection often improves social welfare (National Academy of Science 1965). New food safety research and data collection can help to reduce uncertainty in decision making and reduce costs due to the misallocation of resources. For convenience, we distinguish data needs into two broad categories: physical/biological science data needs and economic/behavioral science data needs.

Physical/biological data are needed to quantify the underlying technical relationship in the production, processing, and marketing of agricultural commodities. In particular, this type of data is needed to establish how chemical and microbial contamination levels relate to production, processing, marketing, product quality, and health risks. Critical to the ability to obtain such information is the availability of tests for detecting chemical residues and microbial levels in an economically efficient way.

Economic/behavioral data are necessary to understand factors affecting the choices of market participants. To alter food production and handling practices to reduce foodborne risks, one must know something about the economic tradeoffs facing producers, processors, and consumers. For example, reducing chemical use on farms and in the marketing system requires knowing the impacts of pesticide, fuel, and la-

bor prices and alternative pest control strategies (such as alternative chemical pesticides, biological controls, product quality, crop rotation, weeding, and integrated pest management) on pesticide usage, productivity, and costs. Both physical/biological and economic/behavioral data have important economic implications for foodborne risk assessment and risk management. Decision makers, whether they be consumers, producers, or regulators, need to evaluate the economic consequences of alternative choices, and these choices depend in part on the underlying physical and biological relationships. Especially important is the recognition and measurement of risk "tradeoffs," i.e., how does the reduction of one source of risk affect exposure to other types of risk and what are the economic costs of this tradeoff?

Physical/Biological Data Needs

New data are needed to pinpoint the food safety hazards associated with the greatest social costs, more definitively identify the foods which are vehicles for these hazards, clarify where the hazardous chemicals or microbes are entering the food chain, and clarify the underlying physical and biological relationships of possible control programs.

Microbial risk assessment data are needed to better estimate the extent and range of human illnesses from microbially contaminated food because there is no consensus on how many cases of foodborne illness occur annually. Yearly, 400–500 foodborne outbreaks involving a total of 20,000 to 25,000 persons are reported to the Centers for Disease Control (CDC).² Recently CDC researchers published their "best estimate" of acute cases of foodborne microbial illness—6.5 million cases annually in the United States, of which 9,000 end in death (Bennett et al.). Researchers in the Food and Drug Administration have examined human illness data and estimated 33 million cases of foodborne illness annually (Garthright, Archer, and Kvenberg). None of these three estimates is universally accepted, and all omit some classes of acute illness (for example, most foodborne viruses lack identifying tests, although viruses are thought to be an important cause of foodborne illness). Also, complications and chronic conditions following

² About 60% of the organisms causing outbreaks are not identified (Centers for Disease Control, Bean et al.). The uncertainties in estimating microbial risks come from underreporting acute illnesses as well as complications following acute illness.

the acute foodborne illness, such as septicemia, arthritis, or cardiac dysfunction, may be difficult to associate with their foodborne cause (Archer 1984, 1985; Mossel).

The social costs of foodborne illness, using the cost of illness approach to quantify medical and productivity losses, have been estimated in the literature at \$4 billion to \$8 billion annually (Garthright, Archer, and Kvenberg; Roberts; Todd). The social costs of complications or chronic conditions following the acute infection may be more costly than the acute illness. For example, in one survey people with arthritis were willing to pay 22% of their household income to get rid of arthritis (Thompson). Roberts and Frenkel estimated that the acute illness costs of fetal exposure to *Toxoplasma gondii*, a common contaminant of pork, were dwarfed by the lifetime productivity losses, additional training costs,

and increased living costs for persons developing mental retardation.

Given advances in epidemiology and rapid tests identifying microbial pathogens, a superior tracking system for microbial risks could be established for use by regulators. The Centers for Disease Control now collects some of this data in a fragmentary fashion in its foodborne disease outbreak reports. Expanding the scope of CDC's efforts, linking it to data collected in other health surveys, and linking it more closely to food vehicles would improve the ability of FDA and FSIS to quantify foodborne disease risks and fashion regulatory programs to reduce human illnesses from foodborne sources (table 1). Particular attention should be given to strengthening identification of microbial foodborne deaths, both because of the importance of achieving a consensus on the number of foodborne deaths

Table 1. Data Needed for Estimating Microbial Health Risks

Data Need	Possible Solution	Likely Cost
Better recording of existing data	<ul style="list-style-type: none"> Expand CDC's passive laboratory-based reporting for 9 foodborne microbial pathogens to the 30 with tests. (Expanding the use of tests by state laboratories would increase costs more.) Expand CDC's laboratory-based reporting to make surveillance active in selected counties (sentinel counties). 	Around \$500,000/year Additional \$1 million/year Around \$75,000
Better demographic data on who becomes ill	<ul style="list-style-type: none"> Study various methods of making reporting active and estimate likely costs and benefits of increased identification of cases. 	
Increased knowledge about risk factors	<ul style="list-style-type: none"> Study individual susceptibility as well as food handling and consumption practices for people becoming ill and control groups who do not become ill. 	Around \$500,000/pathogen
Increased knowledge of disease severities	<ul style="list-style-type: none"> Increase the number of pathogens investigated in depth with surveys of people becoming ill from specific foodborne pathogens. 	Each survey may cost around \$100,000–\$200,000
Increased knowledge on deaths from foodborne sources	<ul style="list-style-type: none"> Expand the sentinel county survey to include all counties in the United States for a few selected pathogens to develop a solid baseline on foodborne deaths. Study improving the identification of foodborne deaths in the National Death Index by improving entry of laboratory test data, by revising the death certificate to specifically ask for laboratory data, etc. 	Around \$1 million/pathogen for passive surveillance, and much more for active surveillance Around \$300,000
Increased knowledge about hospitalizations caused by foodborne pathogens	<ul style="list-style-type: none"> Study improving identification of foodborne pathogens in the National Hospital Discharge Survey (NHDS) by 1) examining a sample of the 150,000 hospitalizations under ICD-9 "008-unspecified intestinal infections" to see if medical records contain laboratory data, 2) examining septicemia/bacteremia records for 1 year to determine if foodborne pathogens are involved, 3) estimating the cost of revising the face sheet for entry of NHDS data to include laboratory data and where sample taken (blood, stool, cerebrospinal fluid), and 4) comparing state systems. 	Unknown, but likely to be less than \$200,000/study
Quantification of chronic conditions caused by foodborne pathogens	<ul style="list-style-type: none"> Review the literature and estimate the likelihood that foodborne pathogens are important in certain chronic disease syndromes such as reactive arthritis, neurological disorders, cardiac dysfunction, food allergies, gastritis, kidney and liver dysfunction, etc. 	Literature review around \$100,000; original research more costly

and because of the high social costs associated with deaths.

Chemical risk assessment data are needed to identify and clarify the range and extent of human health risks. Assessment of chemical risks are influenced by four major factors: chemical usage levels, residue concentrations, potency/toxicity, and consumption levels. To compensate for limited data, scientists often make different assumptions in their analyses. Limited data can lead to dramatically different conclusions regarding the absolute risk levels and the relative ranking of sources of risk. Results from two studies are shown in table 2. The National Academy of Sciences (NAS) study estimates, using EPA risk assessment procedures, are often many orders of magnitude higher than the Archibald and Winter study. For example, the NAS study estimated that chlordimeform used on tomatoes increases cancer risk by 479 additional cases per million people over their lifetime. In

contrast, Archibald and Winter, using FDA residue data to indicate human exposure, estimate no increase in cancer risk from its use on tomatoes. The risk assessments in these two studies suggest vastly different regulatory concerns and strategies. Better and more standardized data and procedures are needed to clarify the actual risks involved.

EPA studies have not developed a consensus on foodborne disease risks. In 1987, the agency estimated the chemical residues on food caused no more than 6,000 excess cancer cases annually, a rank of third out of fifteen risks studied (EPA 1987). A more recent outside advisory panel study estimated the risks of foodborne chemical contamination to rank twenty-fifth out of thirty risks evaluated (EPA 1990).

Data needs for chemical risk assessment include more tests to assess the toxicity of chemicals. The NAS study used only twenty-eight of fifty-three suspected oncogenic pesticides be-

Table 2. Estimated Cancer Risk for Selected Foods

Commodity/Pesticide	NAS (EPA Tolerance)	Archibald & Winter (FDA Residue Data)
(excess cancers per million lifetimes)		
Tomato		
Acephate	14	0.0017
Captafol	191	0.0033
Chlordimeform	479	0
Chlorothalonil	61	0.23
Permethrin	31	0.088
Other	83	0.007
Total	859	0.33
Lettuce		
Acephate	16	0.025
Captan	55	0.011
Folpet	42	0.054
Permethrin	143	0.8
Other	5	0
Total	261	0.89
Apple		
Captan	28	0.07
Chlordimeform	1,368	0
Folpet	42	0
Other	24	0
Total	1,462	0.07

Source: Data are from Archibald and Winter.

Note: Because of limited availability of data, NAS assumed a worst case scenario that each pesticide was present on the approved commodities at the tolerance level. The NAS assumption implies that each pesticide was applied to every acre of every crop for which the chemical was approved and was used in its most intensive manner approved. While this assumption is useful in obtaining an upper bound on usage, it is unlikely to represent actual usage because application rates depend on a number of environmental and economic factors. In addition, some of the pesticides are substitutes for one another, making it unlikely that they would be used jointly at their maximum levels, and some of the older pesticides have been replaced by cheaper and/or more effective pesticides. For example, Carlson points out that linuron, an older pesticide, is approved for use on wheat, corn, and other crops, but surveys indicate little use except for soybeans. The Archibald and Winter study assumed residue levels were equal to actual residue levels found in FDA surveys of supermarket foods.

cause of lack of tests to obtain the potency factors (toxicity) for the remaining twenty-five compounds. While these twenty-eight compounds comprise some of the most widely used pesticides, toxicity data on the remaining twenty-five compounds are necessary to conduct a complete analysis.

Improved models of toxicity also are needed (Finkel, Ames and Gold). Toxicity is extrapolated from animal feeding studies in which sensitive rodent species are fed large doses of the chemical. The dose-response relationship is usually assumed to be linear throughout the entire dosage spectrum. In some cases, multistage dose response models are used. These models extend the simple linear model to include a series of two or more linear segments. When extrapolating from rodents to humans, adjustments are made for lifespan, bodyweight, and species sensitivity. Data needs include improved dose-response models, information on the health effects of metabolites of chemicals as well as the chemical itself, and an expanded scope of health outcomes to include not only tumors but also neurological effects, effects on fetuses, effects on the functioning of the immune system, and other adverse health effects.

Improved data on levels of microbial or chemical hazards in foods and whether these levels are a risk to human health are essential for designing effective control programs. Identifying which foods contain contaminants are suggestive, but the link to human illness is crucial because of variability in the dose needed to cause illness and impact of food preparation and handling methods which often dramatically reduce contamination levels (or in the case of bacteria, can also increase contamination levels by cross-contaminating clean foods or bacterial multiplication).

More data are needed to identify populations particularly vulnerable to foodborne illness and identify their likely disease severities. Some people may be more vulnerable either because of diet or individual characteristics. Depending on the microbial pathogen, such groups can include infants, persons taking antacids or antibiotics, persons who are genetically low acid secretors, persons with new organ transplants or on chemotherapy, diabetics, alcoholics, AIDS patients, and the elderly with no underlying disease. For chemical contaminants, the young are particularly vulnerable because of their long life expectancy, higher metabolism, and perhaps greater intake per unit of body weight. Identifying

groups susceptible to foodborne disease requires the linkage of usual food intake patterns with actual contamination levels.³

Funding for the collection of new data is essential for developing a statistical framework for estimating which foods are vehicles for various chemical or microbial contaminants causing human illness. For chemicals, the new USDA data initiative is collecting much of this data (see below). For microbial risks, data needed include:

(a) Increased funding for CDC investigations of foodborne disease outbreaks to include, routinely, the replication of likely food-handling procedures which could have caused the outbreak. This exercise would improve the identification of the causative organisms (currently 60% are never identified) and would yield more specific information on what conditions are likely and unlikely to lead to foodborne disease outbreaks for specific pathogens—information crucial for setting up HACCP (Hazard Analysis at Critical Control Points) systems.

(b) Funding to assist states, counties, and cities in identifying and investigating foodborne disease outbreaks and perhaps funding for CDC to set up new data collection systems from states.

(c) Funding of additional case/control studies to determine the etiological fraction, e.g., identify what foods are associated with which foodborne pathogens. While these studies are expensive,⁴ they may be the most cost-effective method of building a statistically valid data base for newly identified foodborne pathogens and developing a consensus among public health professionals, industry, and regulators as to the relative importance of various foods as vehicles for specific microbial foodborne diseases.

Identification of the entry point where hazards enter the food chain is also missing data. Once the foods containing contaminants posing human health risks have been identified, the point of entry into the food chain needs to be identified. Because microbial and chemical contaminants can enter the food chain at any point from the

³ The NAS study of pesticides assumed that food consumption was at the 95% confidence limit for food weight for each of 273 food forms in the 1977–78 USDA Nationwide Food Consumption Survey. While this may understate exposure for certain individuals, it overstates average exposure and does not account for substitutions in consumption that are likely to occur. For example, an individual consuming large quantities of one item is likely to consume less of another during a particular day. In other words, the distribution of usual intake is likely to show less variation than that observed in a one-day intake survey and hence have a lower 95% confidence limit.

⁴ The Seattle-King County study identifying chicken as the greatest source of campylobacteriosis risk cost FDA \$800,000.

farm to the point of consumption, control options exist at all these points. This would involve sampling at all levels of the production and marketing system to develop a statistically valid data base for both microbial and chemical risks. Sampling should start with the foodborne disease hazards causing the highest social costs and those that can be easily controlled.

Production function tradeoffs affecting microbial and chemical contamination of foods are needed for farm production, processing, and food-handling practices. Some of the production function data are known, but much remains unknown. For chemical hazards, basic data are needed on types and amounts of chemicals used on specific crops by geographic regions, crop yields and quality at alternative chemical usage levels, alternative input substitutions, and cropping patterns. For example, what are the effects of integrated pest management on specific crop yields and quality in specific geographic regions, under specific weather conditions, and what is the impact on agricultural chemical usage levels and residues in foods.

For microbial hazards, data are needed identifying likely microbial contamination levels with various farm management procedures, processing procedures, transportation practices, and food handling procedures (Wenger et al.). For instance, some scientists believe that *Salmonella* and *Campylobacter* can be controlled or eliminated in poultry flocks by control of eggs, feed, and environmental conditions (such as chlorinating drinking water). Economic data are needed for evaluation of these alternatives.

Economic and Behavioral Science Data Needs

Private and public risk management strategies require that the economic and other behavioral factors that affect decisions of suppliers, consumers, and regulators be taken into account (table 3). Options for control programs need to be evaluated to determine the combination that maximizes social welfare. The National Academy of Sciences has concluded that the current meat and poultry inspection system based on visual inspection cannot detect the most important microbial risks (1985, 1987, 1990a,b). The changing scientific ability to rapidly test for microbial pathogens creates an opportunity to revamp regulatory approaches. For example, FSIS is using recent advances in rapid detection methods for a *Listeria* control program.

Table 3. Economic Data Needs

<u>Economic Assessment of Risks</u>
Identification of hazard
Toxicity/microbial virulence
Exposure
Costs: consumers, industry, public
<u>Demand Issues</u>
Demand for and value of information
Risk perception
Demand for product attributes (appearance/taste/safety tradeoffs)
Willingness to pay/contingent evaluation
Consumption and exposure
Consumer demand and intercommodity relationships
<u>Private Strategies and Supply Issues</u>
Alternative production/management strategies (HACCP)
Costs and benefits of risk control/management strategies
Vertical market integration/contract specification for inputs
Industry self-regulation
Third-party certification/labelling/branding
Cost of production
<u>Market Interrelationships and Government Intervention</u>
Supply/demand
Farm program and input (agrichemical) usage
Marketing orders/grades and standards/agrichemical usage
Inspection/labelling/standards for microbial and chemical residues
Environmental effects
Federal/state/foreign jurisdiction
Tracking systems for feed, food, and chemicals

(a) Estimate the benefits and costs of farm management strategies to prevent farm-level microbial and chemical contaminants. Basic economic data are needed to estimate the possible cost increases associated with farm management strategies identified in the physical/biological data needs section. One such farm-level analysis would be estimating costs of eliminating parasites, such as *Taxoplasma gondii* and *Trichinella spiralis*, on hog farms. New data from the National Animal Health Monitoring System may be helpful in answering questions about the effectiveness and cost of alternative farm management strategies for reducing microbial contamination of food animals. In estimating the human illness prevention benefits of farm control, it is advisable to estimate both the cost of illness avoided as well as the willingness to pay for safer meat and poultry products. (The cost-of-illness measure is more easily understood by the public and members of Congress, while the willingness-to-pay measures are more comprehensive but are still in the developmental stage of public acceptability.)

(b) Predict/understand how farmers choose alternative production practices. What price differential between pesticides will trigger a switch to another chemical or cause a switch to organic production? How do farmers obtain information about farm chemical risks, prices, and effects on pests? What are the resource requirements and substitutions in terms of management skills, labor, capital, and variable inputs (and their related costs, productivity, and production risks) for alternative farm production systems? How do farmers evaluate their own risks to pesticide application exposure, environmental damage, and ground water contamination and how does this knowledge affect behavior?

(c) Identify the benefits and costs of control strategies beyond the farm gate. For parasites, thorough cooking or irradiation are the only options for reduction of human exposure because parasites do not multiply in the meat and are hardy survivors during traditional food handling, processing, and transporting. For most bacteria, there are other possibilities, because cross-contamination and multiplication are likely during food processing, handling, and transporting. Another class of organisms are bacteria-producing, heat-stable toxins which cannot be eliminated by thorough cooking; here the keys are either preventing bacterial contamination or preventing toxin production during processing, handling, and transporting.

For chemicals, we need to understand processors' choices of fumigants, food additives, and other chemicals. Data on producer and marketer strategies to manage food safety need to be collected to evaluate the efficiency and viability of private market alternatives to government regulation. Recent evidence suggests that food safety is becoming an important marketing strategy for food retailers. Kaufman and Newton note that the use of third-party residue testing and certification services for fresh fruits and vegetables grew from a single supermarket chain in 1987 to 14 retailers in 1989 operating more than 740 grocery stores. Despite its growing importance, little data are available on the costs, prevalence, and types of products included, specific tests or assurance methods, labelling, availability of organic produce, and related market assurance activities. To what extent do marketers enter into forward contracts with farmers that include specifications on pesticide usage and residues? Would marketing grades specifying food safety attributes improve market efficiency?

(d) Study consumer behavior to determine ex-

posure and consumers' food safety preferences. Consumers, to varying degrees, have some control over foodborne disease risk. For example, they can wash produce, peel skins, and remove outer leaves where pesticide residues and microbes tend to concentrate; purchase a variety of foods; avoid more risky foods; buy organic produce or "certified residue free;" cook meat, poultry, and seafood thoroughly; refrigerate cooked foods promptly and keep refrigeration temperatures around 35 degrees Fahrenheit; and practice improved kitchen sanitation. Food safety policies can also affect the prices, qualities, and availability of foods. Thus, there is a need for this basic demand data to evaluate market impacts and consumer welfare.

Data are needed to conduct research on underlying consumer behavior issues, such as consumer demand for alternative product attributes, e.g., flavor, appearance, pesticide residues, and product price, willingness to pay for safer foods, demand for organic and pesticide-free produce. Very little data are currently available to research these issues. This would necessitate the collection of data at the household or individual level. Questions on consume risk perception, awareness, and attitudes also need to be included in survey instruments. If a link between cosmetic appearance and pesticide residues is shown to exist, how willing are consumers to make tradeoffs between cosmetic appearance and pesticide residues? These research issues will require new surveys of consumers with new survey instruments. Conjoint analysis and contingent market evaluation techniques will be necessary to elicit some of the underlying aspects of behavior.

The Alar controversy indicated increased public concern about particularly vulnerable groups. It also suggested that designing different regulatory programs for different segments of the population may optimize social welfare. Special regulatory interventions may be desired for some of these groups, such as special labels on meat and poultry identifying product risks, special labels with handling and cooking instructions, certification programs for products with reduced pathogen levels. (FDA has been targeting AIDS patients with special food handling advice [ref]). Focus group work would be necessary to evaluate how these programs might be valued and received by consumers.

(e) Once technically feasible strategies have been identified, economical and socially optimal regulatory programs and enforcement strategies need to be identified and evaluated. Ac-

tions such as slaughterhouse tests, farm level inspections, fines for contaminated carcasses, or required cooking for carcasses from farms with contaminated animals are some of the regulatory options available for controlling contamination of raw meat and poultry. How each strategy can be expected to alter farm management practices should be modeled and the benefits and costs of each alternative compared. Strategies at later points in processing, handling, and transporting also need to be evaluated. One such study could compare the legal tools, enforcement, and effectiveness of federal programs and policies.

Regulators need better data to trace agricultural products through the marketing system. This information would allow for improved crisis management of foodborne hazards, improved enforcement activities, and provide a mechanism for shifting product liability back to the responsible party.

The USDA Food Safety Data Initiative

The USDA food safety data initiative is a closely coordinated effort of four agencies to gather data on agricultural chemicals in the food system (table 4). For the USDA effort, the National Agricultural Statistics Service (NASS) is collecting data on pesticide use at the farm level, the Agricultural Marketing Service (AMS) is collecting pesticide residue data on fresh fruits and vegetables, the Human Nutrition Information Service (HNIS) is developing exposure models based on its surveys of food intake of individuals, and the Economic Research Service (ERS) is analyzing the economic implication of alternative means of reducing residue levels. The USDA initiative is being coordinated with FDA and EPA, the two other federal agencies with responsibility for regulating pesticides in the food supply.

The National Agricultural Statistics Service

will be responsible for collecting pesticide usage levels and practices. The data will allow for the identification of usage levels for specific chemicals by crop and geographic region. Economic data on farm resources, production practices, and input prices also will be collected. During the first year of the initiative, data collection efforts will focus on chemical usage on domestically produced fresh marketed vegetables in the five major vegetable-producing states. Efforts during the second year will focus on fresh fruits and nuts.

The Agricultural Marketing Service, working with individual states, will be responsible for collecting statistically based data for pesticide residues in selected fresh fruits and vegetables grown in the United States. The data will be used to improve the quality and quantity of information on actual pesticide residues for regulatory purposes as well as to inform and assure the public on actual pesticide residue levels. AMS is working closely with EPA in prioritizing chemicals for residue testing and deciding upon appropriate laboratory procedures. AMS will provide funds and other resources to participating states to adapt their programs to a standardized system for sample collection, analysis, and reportage.

The Human Nutrition Information Service will develop a residue exposure assessment system. HNIS survey data on individual intake will be linked to residue data obtained from AMS data collection efforts. By linking HNIS survey data on individual intake to residue data from AMS data collection efforts, the system will provide data on the potential exposure of individuals to pesticide residues found on agricultural products. In addition, data on the amounts of more than 5 000 different food items will be translated back to the agricultural products of which they are comprised via recipes for linkage to the farm pesticide usage data.

Discussion

The efficient control and management of food safety risks requires the synthesis of data and information from a number of physical, biological, and behavioral sciences. The process of information synthesis is technologically complex and fraught with uncertainty. The uncertainty arises from three major sources: limited science, limited technology, and limited data. One should not expect new data to completely eliminate uncertainty from the food safety risk assessment

Table 4. USDA's Food Safety Initiative

1. Identifies the set of residues to be monitored and their toxicity.
2. Identifies entry point into the food system and surveillance point.
3. Develops sampling requirements, procedures, alternatives, analytical methods of detection, and cost.
4. Measures exposure.
5. Coordinates data-gathering efforts.
6. Collects related economic data affecting the decision-making behavior of market participants: producers, consumers, and regulators (see table 3).

and management problem because of diminishing returns and increasing marginal costs. There is a tradeoff in resource allocation, and one must choose wisely among those activities designed to reduce uncertainty and improve decision making such as the collection of new data. A number of important factors are identified in this paper that provide insight into the factors that influence food safety decisions and the degree of uncertainty regarding their impact. This information, together with costs of obtaining data to improve our knowledge, will guide us in the selection of new food safety data collection efforts.

A science-based approach to food safety issues is data intensive. Data are needed for risk assessment, identification of points of entry into the food chain, tracking substances as they move through the food system, analyzing the behavior and perceptions of market participants, communicating disparities between real and perceived risks, and measuring the costs and benefits of alternative control options. For example, the National Academy of Sciences in 1969 identified steps that could be taken to prevent *Salmonella* contamination of raw meat and poultry. These procedures have not been implemented, and human illness of salmonellosis have increased in the United States. Economic studies of the costs and benefits of alternative farm management practices are needed before industry and regulators will be willing to make decisions about adopting new practices.

In addition, the changing nature of scientific knowledge in the food safety area means that new possibilities for assessing and controlling food safety risks are emerging. Federal funding of basic research to speed up the emergence of scientific knowledge is a standard economic textbooks' example of a public good. With a public good, the private sector cannot capture the full return to society of the new scientific knowledge; once the invention has occurred, anyone can use the knowledge regardless of whether or not they pay for it.

Current estimates of foodborne disease risk suggest that microbial and other natural toxic components are the greatest sources of risk. Current progress on identifying microbial risks and sources in food is pointing the way to control options. However, additional data are needed to conduct benefit/cost analyses of alternative control options and will require interdisciplinary cooperation to combine the underlying physical and technical relationships together with the economic and behavioral aspects of market par-

ticipants. In conclusion, optimal public and private strategies to assess and manage food safety issues require that additional data be collected, analyzed, and integrated into public and private decision making. The greatest social welfare gains are likely to be garnered by collecting data on microbial hazards and control options—the source of the greatest human health risk and a compliment to USDA's data initiative on chemicals.

References

- Akerlof, G. A. "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism." *Quart. J. Econ.* 83(1970):488–500.
- Ames, B. N., and L. S. Gold. "Too Many Rodent Carcinogens: Mitogenesis Increases Mutagenesis." *Science* 249(1990):970–71.
- Archer, D. L. "Diarrheal Episodes and Diarrheal Disease: Acute Disease with Chronic Implications." *J. Food Protection* 47(1984):321–28.
- . "Enteric Microorganisms in Rheumatoid Diseases: Causative Agents and Possible Mechanisms." *J. Food Protection* 48(1985):538–45.
- Archer, D. L., and J. E. Kvenberg. "Incidence and Cost of Foodborne Diarrheal Disease in the United States." *J. Food Protection* 48(1985):887–94.
- Archibald, S. O., and C. Winter. "Pesticides in Our Foods: Assessing the Risks." *Chemicals in the Human Food Chain*, ed. Carl K. Winter, James N. Seiber, and Carole F. Nuckton. New York: Van Holt Reinhold, 1990.
- Bean, N. H., P. M. Griffin, J. S. Goulding, and C. B. Ivey. "Foodborne Disease Outbreaks, 5-Year Summary, 1983–1987." *J. Food Protection* 53(1990):711–28.
- Bennett, J. V., S. D. Holmberg, M. F. Rogers, and S. L. Solomon. "Infectious and Parasitic Diseases." *Closing the Gap: The Burden of Unnecessary Illness*, ed. R. W. Amler and H. B. Dull, pp. 102–14. New York: Oxford University Press, 1987.
- Carlson, G. A. "Risk Assessment and Regulatory Priorities for Pesticide Residues in Food." *Pesticide Residues and Food Safety*, ed. W. Ferguson and P. Smezdra, pp. 32–39. Washington DC: U.S. Department of Agriculture, Econ. Res. Serv., AER Staff Rep. No. AGES 9110, 1991.
- Centers for Disease Control. "Foodborne Disease Outbreaks, 5-year Summary, 1983–1987." *MMWR* 39(1990):15–59.
- Environmental Protection Agency. *The Report of the Human Health Subcommittee, Appendix B*, EPA SAB-EC-90-021B. Washington DC, 1990.
- . *Unfinished Business: A Comparative Assessment of Environmental Problems*. Washington DC, 1987.
- Experiment Station Committee on Organization Policy. *Research Agenda for the 1990s*, ed. Terry G. Huff. Texas Agricultural Experiment Station, Feb. 1990.
- Farley, D. "Food Safety Crucial for People with Lowered Immunity." *FDA Consumer*, July 1990, pp. 7–9.

- Finkel, A. M. *Confronting Uncertainty in Risk Management: A Guide for Decision Makers*. Washington DC: Resources for the Future, 1990.
- Garthright, W. E., D. L. Archer, and J. E. Kvenberg. "Estimates of Incidence and Cost of Intestinal Infectious Diseases in the United States." *Public Health Rep.* 103(1988):107-16.
- Hushak, L. J., W. S. Chern, and L. Tweeten. "Priorities for Data on Agriculture and Rural Areas: A Survey of Agricultural and Rural Scientists." A Project of AAEE's Economic Statistics Committee, draft, July 1989.
- Kaufman, P., and D. Newton. "Retailers Explore Food Safety Assurance Options." *Nat. Food Rev.* 13(1990): 11-15.
- Mossel, D. A. "Impact of Foodborne Pathogens on Today's World, and Prospects for Management." *Animal and Human Health* 1(1988):13-23.
- National Academy of Sciences. *An Evaluation of the Salmonella Problem*. Washington DC: National Academy Press, 1969.
- . *Basic Research and National Goals*. Report to the Committee on Science and Astronautics of the United States House of Representatives. Washington DC, 1965.
- . *Cattle Inspection*. Washington DC: National Academy Press, 1990.
- . *Meat and Poultry Inspection: The Scientific Basis of the Nation's Program*. Washington DC: National Academy Press, 1985.
- . *Poultry Inspection: the Basis for a Risk-Assessment Approach*. Washington DC: National Academy Press, 1987.
- . *Regulating Pesticides in Food: The Delaney Paradigm*. Washington DC: National Academy Press, 1987.
- . *Risk-Assessment in the Federal Government: Managing the Process*. Washington DC: National Academy Press, 1983.
- Office of Management and Budget. *Regulatory Program of the United States Government: April 1, 1990-March 31, 1991*. Washington DC: Executive Office of the President, 1990.
- Roberts, T. "Human Illness Costs of Foodborne Bacteria." *Amer. J. Agr. Econ.* 71(1989):468-74.
- Roberts, T., and J. K. Frenkel. "Estimating Income Losses and Other Preventable Costs Caused by Congenital Toxoplasmosis in People in the United States." *JAVMA* 196(1990):249-56.
- Thompson, M. S. "Willingness to Pay and Accept Risks to Cure Chronic Disease." *Amer. J. Public Health* 78(1986):392-96.
- Todd, E. C. D. "Preliminary Estimates of Costs of Foodborne Disease in the United States." *J. Food Protection* 52(1989):595-601.
- U.S. Department of Agriculture. *Food News for Consumers: Is Someone You Know at Risk for Foodborne Illness?*, pp. 4-10. Washington DC, spring 1990.
- Viscusi, W. K. "Toward a Diminished Role for Tort Liability: Social Insurance, Government Regulation, and Contemporary Risks to Health and Safety." *Yale J. Regulation* 6(1989):65-107.
- Wenger, J. D., et al. "Listeria monocytogenes Contamination of Turkey Franks: Evaluation of a Production Facility." *J. Food Protection* 53(1990):1015-19.

Data Needs to Assess Environmental Quality Issues Related to Agricultural and Rural Areas: Discussion

David E. Ervin

Environmental quality issues are now prominent in virtually all comprehensive agricultural and rural policy discussions. The demand for data to analyze these issues, therefore, is strong, reflecting public concern about agriculture's diverse impacts on the environment. Moreover, the current era's environmental issues require new temporal and spatial data because of the prominent roles of uncertainty (e.g., groundwater quality, food safety) and broader geographic coverage (e.g., global change). It is within this context of volatile policy discussions and uncertain science that critical data choices must be made.

Fletcher and Phipps (FP) provide a comprehensive survey of the wide variety of data management issues related to agriculture-environment topics. They do a good job at covering a broad array of questions that are crucial to the construction of a high quality database to support sound science and policy development. In this brief discussion, I will emphasize five points that are consistent with the FP paper and that are especially relevant to aggregate policy analysis.

Monitoring and research questions should drive data collection. The authors make this point, but it merits further emphasis. I am using monitoring in the broadest sense here to include the potential pollutant loadings, environmental performance variables, and demand characteristics for environmental services. While there will always be discussion as to which are the relevant questions in science and policy circles, the

"right" data cannot be collected without first asking the "right" questions. To that end, I would have preferred that FP attempt to specify some of what they feel are the dominant questions regarding agriculture and environmental quality. An example from the Just and Antle analysis is, to what extent a farmer input behavior is causally related to by natural resource-environmental characteristics.

Different levels of resolution are needed for general monitoring and for analytical research. Monitoring information can be very useful in defining relevant research questions, and research findings may identify missing environmental monitoring indicators. While these interrelationships exist, an agriculture-environmental quality database strategy should recognize that the two activities often require different types of data. For example, national and state agricultural chemical use data can be helpful for targeting research efforts to those geographic areas and/or production systems more likely to experience environmental problems.

Existing data and science are inadequate to provide the theoretical framework and sampling procedures to understand aggregate agriculture-environment interactions. Thus, aggregate data management involves many choices under very uncertain conditions. A good example of this point is an exercise recently attempted under the president's Water Quality Initiative in agriculture. The objective was to devise a matrix of major agricultural activities and groundwater vulnerability conditions to sample geographic areas. Not only were data unavailable to assign cell probabilities, but the cooperating physical scientists indicated that an acceptable vulnerability measure is not at hand. However,

David E. Ervin is chief, Resource Policy Branch, Resources and Technology Division, Economic Research Service, U.S. Department of Agriculture.

The comments and views expressed in this article are the author's and not necessarily those of the U.S. Department of Agriculture.

by asking a policy-relevant question, future data collection can be designed to reduce that uncertainty and improve scientific understanding. How do we answer aggregate water quality economics questions in the meantime? Perhaps the best approach is to use smaller area study knowledge to parameterize input-input, input-output, and output-output relationships in aggregate regional/national simulation models and then use the price-quantity endogenous aggregate results to analyze economic and environmental adjustments in what are considered the "major" agriculture-water quality cases.

Emphasize the understanding of economic behavior related to natural resource/environmental conditions. Just and Antle have presented a very appealing conceptual framework with which to analyze agriculture-environmental quality interactions. Their emphasis is on the joint spatial and temporal distributions of physical and economic data for analyzing the effects of alternative policies. Underlying their discussion is the fundamental point that we must understand economic behavioral processes in relation to the environment to predict the consequences of those policies. With the appearance of a plethora of intriguing environmental process simulation models, there is a distinct risk of economists not allocating enough effort to the economic behavioral questions. Clearly, our comparative advantage is applied economic, but incorporating environmental process detail is needed to un-

derstand behavior. The data design implications are straightforward.

Geographic information systems technology holds considerable promise for overview monitoring, but its analytical research potential is unclear. As FP indicate, GIS is a powerful emerging technology. Its strength lies in organizing different and sometimes disparate data sets to portray spatial economic and environmental relationships. The best level of resolution for GIS use is still unclear, i.e., small area studies or aggregate overviews. The answer is probably not uniform but depends upon the economic and environmental processes under analysis. If close detail is necessary to understand variation in significant economic and environmental patterns (e.g., input use and ag chemical leaching), then small area GIS applications are warranted. In contrast, if those relationships are fairly constant over large spatial units, then aggregate summarization is possible. What is not at all clear is whether GIS technology can be effectively used in behavioral analytic exercises. To date, such analyses are not forthcoming. Perhaps we are just in the early stages of technology adoption.

Reference

- Just, Richard E., and John M. Antle. "Interactions between Agricultural and Environmental Policies: A Conceptual Framework." *Amer. Econ. Rev.* 80(1990):197-212.

Data Needs to Address Economic Issues in Food Safety: Discussion

Julie A. Caswell

What are the economically significant food safety risks facing the U.S. population? What are the most cost-effective means of controlling (managing) these risks? These are straightforward questions; yet, as the paper by Roberts and Smallwood indicates, the answers require a complex array of information only parts of which are currently available. Discussion of data needs necessarily generates a list. But what makes for a good list, as opposed to a laundry list?

Roberts and Smallwood organize their list by dividing food safety data needs into two broad categories: physical/biological and economic/behavioral. They also consistently distinguish between food-related risk from microbial and chemical sources. The latter includes possible risks associated with pesticides, animal drugs, environmental contaminants, and additives. In the physical/biological area, their list of data needs includes risk assessment information for risks from both microbial and chemical sources in food products, information on levels of contaminants, risks to human health, entry points, and production function tradeoffs. These correspond roughly to the first three steps of the National Academy of Sciences' (NAS) four-step risk assessment procedure: hazard identification, dose-response assessment, and exposure assessment. The economic/behavioral category of data needs corresponds to the fourth step in the NAS's process: risk characterization. It is helpful, I think, to locate individual data needs in the context of this existing NAS framework and to piggyback off prior efforts. The Roberts and Smallwood paper provides a very serviceable list of data needs from which to carry forward this process.

Economists are ultimately end users of the array of physical/biological data needed for analysis of food safety issues. Their major task in this area is to work with researchers from other disciplines to insure that data collection systems are designed to generate information in formats

amenable to economic analysis. Because these data sets are destined to be put to multiple analytical purposes, quality of design is critical. An example of this involvement in survey design is contained in recent work by Carriquiry, Jensen, and Nusser. They suggest design features for the USDA Nationwide Food Consumption Survey to insure that statistically sound measures of distributions of usual daily exposures for individuals can be generated. This type of work is very important to the availability of useful data sets for the 1990s and beyond.

In the economic/behavioral category of data needs, economists should, of course, be providing leadership in data collection efforts. The table of categorical data needs in this area developed by Roberts and Smallwood is particularly helpful in emphasizing the set of private (consumer, firm) strategies and public (government) policies that affect the benefits and costs, and the distribution of benefits and costs, associated with management of food safety risks. Of great interest here are recent shifts in firms' incentives to produce, label, and advertise "safe" products as a result of increased consumer demand for safety attributes in food products (see e.g., Ippolito and Mathios). This new market-based incentive structure will interact with the federal government's regulatory system in ways that promise to be difficult to evaluate from a social benefit/cost perspective.

In the end, economists play an integrative role bringing together the physical/biological and economic/behavioral data to answer the questions I began with: What are the economically significant food safety risks and what are the most cost-effective means of managing them? An excellent road map to this type of effort can be found in recent work by Curtin and Krystynak on salmonella control in Canada. This work integrates data on the incidence of salmonella contamination at different stages in the processing and distribution of poultry products with estimates of the benefits and costs of control strategies at those stages.

An important issue not addressed by the Rob-

Julie A. Caswell is an associate professor, Department of Resource Economics, University of Massachusetts at Amherst.

erts and Smallwood paper is the relationship between the data needed for the management of microbial and chemical food safety risks and that needed for management of nutrition- and diet-related food risks (e.g., heart disease, colon cancer). Yet, within the hierarchy of the federal government's regulatory agenda, all food-related risks will, at least implicitly, be considered as a group and resources will be allocated within this grouping. Fortunately, many data bases, such as the USDA Nationwide Food Consumption Survey, provide information useful for calculating exposure to multiple risks. Elsewhere, however, data needs are not overlapping, and choices about where to commit resources must be made.

What are the priorities among the food-safety-related data needs on Roberts and Smallwood's list? How do politics and public perceptions affect the extent to which these data needs are being addressed? As economists have come to know, and in part respect, there are differences in the food safety priorities of consumers, consumer advocates, processors, retailers, government regulators, and "scientists." If, as Roberts and Smallwood argue, and I agree, social welfare gains will be realized by developing improved food safety information, then the priority-setting

process will be key to the timing and size of these gains. Now that we have the makings of a good list of data needs, our attention should shift to mapping where to start, making maximum use of available data and developing new data sources such as the USDA Food Safety Data Initiative. It is also important, while drawing this map of the food safety neighborhood, to remember that there is a bigger world of consumer risks out there and that the larger scale tradeoffs between management of different classes of risk is the ultimate economic issue.

References

- Carriquiry, A. L., Helen H. Jensen, and S. M. Nusser. "Modeling Chronic Versus Acute Human Health Risk from Contaminants in Foods." *Economics of Food Safety*, ed. Julie A. Caswell. New York: Elsevier Science Publishing Co., forthcoming.
- Curtin, Leo, and Ronald Krystynak. "An Economic Framework for Assessing Foodborne Disease Control Strategies with an Application to Salmonella in Poultry." *Economics of Food Safety*, ed. Julie A. Caswell. New York: Elsevier Science Publishing Co., forthcoming.
- Ippolito, Pauline M., and Alan D. Mathios. "Information, Advertising, and Health: A Study of the Cereal Market." *RAND J. Econ.* 21(1990):459-80.

The Role of Aid and Capital Flows in Economic Development

Uma Lele and Ijaz Nabi

Prior to the debt crisis which commenced in 1982 when Mexico defaulted on its debt repayments, annual capital flows to developing countries amounted to \$100 billion and official concessional aid flows consisted of an additional \$50 billion. After the debt crisis, commercial flows dried up almost completely with net transfers of capital from borrowing countries to their lenders. The concessional flows which had grown dramatically in the 1970s also stagnated in real terms. Over the four post-World War II decades, billions of dollars have been transferred from developed to developing countries. During this period there have been major structural changes in the patterns of external assistance including the graduation of a number of recipients, such as Greece, Taiwan, and South Korea, the decline of the U.S. role as a donor, the rise of Europe and Japan in external assistance, and a shift of aid resources from Asia to Africa. Whether these transfers have succeeded in stimulating growth, reducing poverty, and creating human and institutional capacity to carry out the development effort is a matter of considerable controversy. Econometric studies that have attempted to establish the effect of capital flows on savings, investment, or growth have turned out to show little relationship.

The authors of this paper assembled a group of scholars from developed and developing countries with hands-on experience in development policy to explore the role of external capital in economic development around a well-defined comparative framework. Eleven country case studies on Asian, Latin American, and

African countries¹ were undertaken by these scholars together with four cross-country, issues-oriented studies on bilateral and multilateral aid, food aid, on the process of transition from concessional to commercial capital flows, and on some of the latest instruments of capital transfers. Even while being a small fraction, external finance can make an important contribution to the development process by transferring technology and information, creating institutions, developing human capital, and increasing the supply of critical public goods the shortage of which hamstrings development at early stages. The study addressed the extent to which, and the conditions under which, these potentially positive effects of aid were realized and the factors explaining them.

The overall results of the study are presented first in the form of general lessons for recipients, official donors, and lenders. These are followed by country-specified highlights from which the general conclusions are derived.

A large amount of external capital has not been the most important factor in economic development. Country policies made the most difference. In countries with sound economic policies, a limited amount of well-conceived external assistance helped improve the capacity of the recipient countries to plan and implement overall policies and programs as well as to develop the critical physical and social infrastructure needed to accelerate growth. Even large amounts of external capital could not effectively substitute for unsound internal economic management (i.e., a public investment program not geared to the broad-based provision of physical and human capital, large fiscal deficits, high inflation rates, and low internal mobilization of capital). Outward-oriented development policies involving relatively few distortions and restrictions in

The authors are, respectively, Graduate Research Professor of International Economic Development at the University of Florida, and a senior economist at the World Bank.

The views in this paper are those of the authors and do not necessarily represent the views of the organizations they work for.

This paper is based on the findings of a cross-country study of aid and development coordinated and edited by the authors and published in a book, *Transitions in Development: The Role of Aid and Commercial Flows*. San Francisco: ICS Press, Apr. 1991.

¹ Countries include India, Pakistan, Sri Lanka, Thailand, Egypt, Mexico, Brazil, Columbia, Kenya, Tanzania, and Senegal.

the economy were more conducive to rapid growth compared to inward-oriented policies. When accompanied by a broad-based public investment program in support of education, health, and other social sectors, outward-oriented policies also achieved more equity than inward-oriented policies as they achieve rapid growth in employment. Because removal of distortions through policy reforms is a painful process challenging many vested interests, external capital was able to provide an important cushion to soften the adjustment process until the economies could respond to the reforms. The lack of access to external finance on a long-term assured basis often terminated the reform process which was underway.

Countries that succeeded in achieving a broad-based development of their agricultural sectors achieved more efficient growth than those that neglected their agriculture. Given that agriculture contained the bulk of the population, broad-based agricultural growth played an important role in generating employment and reducing the proportion of the population living in poverty. In several cases, especially in Asia, external financial and technical assistance played an important role in fueling agricultural growth through investment in human capital, institutions, and technological change. Food aid, too, played an important role in Asia in ensuring an adequate amount of food to the poor and vulnerable groups when countries could ill afford to import food. Moreover, food aid provided the crucial foreign exchange needed by countries at early stages of development and frequently facing a balance-of-payments crunch because laying foundations of broad-based agricultural development involved substantial gestation lags before acceleration in the growth of agriculture could occur. Food aid did not play such a positive long-run role in generating agricultural growth in Africa, where internal price and other policy distortions turned out to be insurmountable, although it increased the food supply in the short run.

Another key to broad-based growth was the investment in human capital. Access to universal primary and secondary education, health care, and water were essential to promote broad-based participation of the poor in the growth process. While external aid can play an important role in such investments which tend to face long gestation lags in producing returns, aid did not always play this role. Not only the governments, but even the givers of large aid, neglected the importance of human capital.

Whereas the content of domestic policies explain part of the puzzle of who succeeded in development, the type of aid given also made a big difference to its productivity. Far too much of the aid was given for strategic, commercial, or short-term political reasons and too little for long-term economic development. Aid given for political reasons also tended to be highly unstable. Large fluctuations in the amounts of external capital made it impossible for the recipients to plan uses of aid for long-term investments, e.g., in social expenditures with long gestation lags. Contrary to the permanent income hypothesis, which expects actors to save much of the windfall in income, recipients used the "aid boom" in much the same way as they did commodity booms. They squandered it on unproductive expenditures including growth of governments. The results of the study suggest clearly that, if development is the objective, aid should be given to countries with conducive economic policies, although humanitarian aid may be given to the least developed countries with large concentrations of the poor. Donors also need to avoid fads which make them shift aid allocations from one type of project to another and between project and program aid, although the underlying reasons are well founded for the apparent fashions in aid, e.g., the concern about poverty in the 1970s, macropolicy reform in the first half of the 1980s, and the environment in the second half of the 1980s. The challenge for aid donors has been to address these multidimensional aspects of development successfully.

Aid tying by bilateral donors has been another problem. It has led to transfer of inappropriate technology and economically unjustified projects, leading to suboptimal use of both aid and national funds.

As countries become more economically advanced, they tend to rely less on official aid and increasingly on private sources of capital. Commercial borrowing offers many advantages, but only when it is demanded in moderate quantities; it does not replace internally mobilized resources, and it can be adequately serviced, for example when external terms of trade are favorable. However, the transition from concessional to commercial capital flows is neither smooth nor unidirectional.

Massive amounts of private capital were loaned to developing countries by banks in the 1970s without regard to, or knowledge about, the conditions in which the capital was being deployed or the conditions under which it could be repaid.

Private capital was often used no more productively than public capital, especially politically and commercially driven bilateral aid. It resulted in prestigious megaprojects yielding low returns and an overstretched public sector. Fluctuations in external terms of trade compounded the debt-servicing problem of developing countries saddled with poor investment portfolios. Unless commercial banks invest in developing a better understanding of borrower circumstances and recognize the implications of volatility in external terms of trade for debt repayment, private capital will continue to be poorly allocated. Consequently it will be both unpredictable and unproductive.

In situations of unfavorable policy regimes characterized by serious distortions, external aid and private capital actually proved harmful to countries. By providing massive amounts of capital, donors and lenders contributed to the Dutch Disease by expanding the nontraded public goods sector, exacerbating distortions, and strengthening vested interests. It is clear that commercial banks must share in the burden of this past massive misallocation of capital by sharing in the debt burden of developing countries. Without a major reduction in the debt, it is unlikely that investment in the highly indebted developing countries will resume. On the contrary, those countries who had already made a transition from concessional to commercial capital have had to resort to official assistance to maintain some flow of capital and to reduce the net outward transfers made necessary by heavy debt servicing.

Country performance varied greatly among those selected for analysis. India, one of the poorest countries, received the most concessional assistance in total, although it was minuscule in per capita terms because of its large population. India is a relatively successful case of aid's role in development, especially in the development of its agriculture which led to acceleration in food production and some major reduction in poverty. India needs more concessional aid for investment in the social and physical infrastructure to alleviate its massive poverty. Growth performance of Pakistan, also one of the poorest countries at independence was better than India's, although its social indicators—especially with regard to women—remained some of the lowest. Pakistan experienced major fluctuations in aid because of its strategic geopolitical situation. Wars in Vietnam and Afghanistan determined aid levels. Trade

liberalization in Pakistan was associated with periods of large aid flows and arrested when aid dried up. Egypt, too, received massive aid flows for geopolitical reasons with sources of aid shifting constantly between the Eastern and the Western blocs. Aid contributed to the Dutch Disease. Egypt's economic performance was not commensurate with the levels of aid given to it. Sri Lanka, too, received massive aid and, as in the case of Egypt, sources of aid changed. Eastern bloc countries dominated in aid flows when socialists were in power, and the West dominated when the more market-oriented political party was in power. Aid had little direct influence on Sri Lanka's economic policies however, although socialist governments fared poorly relative to the market-oriented ones, not only in achieving growth but also in fostering equity.

Thailand is one of the most successful cases of rapid growth and broad-based development. The government of Thailand focused on a balanced program of public investment designed to increase the supply of public goods and pursued macroeconomic management that offered an enabling environment for the private sector with a relatively minor role of the government in directly productive activities. Thailand's biggest challenge is to increase internal resource mobilization and maintain its debt servicing under control.

Among the three Latin countries, only Colombia pursued stable and conservative macroeconomic policies, and it was a steady performer. Nevertheless Colombia has encountered difficulties in access to external capital due to the "Latin Syndrome" of private banks. Mexico had a relatively well-managed economy until the 1970s, but the country went on a binge in borrowing at the tail end of the decade and got into macroeconomic difficulties. Mexico has undertaken one of the most ambitious programs of structural adjustment since 1988 and has resumed some access to external capital. Whether this is enough to achieve sustained and broad-based growth remains to be seen. Brazil's miracle was based on an import substitution policy. External capital Brazil borrowed in the 1970s was largely misallocated, and internal reforms have been slower in coming.

Only Kenya performed well among the three African countries selected for analysis. Tanzania and Senegal were large recipients of aid. Their acute import substitution policies were reinforced by external aid, which caused rapid growth of government and reinforced the decline in per

capita incomes caused in part by adverse external terms of trade, and by poor internal economic policies.

The examples of these countries show that external capital is not always a blessing, and too

much of it can be a particularly bad thing. On the other hand, there are several examples in which a small amount of external capital has played an important catalytic role in accelerating economic growth and alleviating poverty.

Investments to Transfer Poultry Production to Developing Countries

Gary Vocke

World chicken meat production is steadily increasing (fig. 1). Much of this meat is produced in controlled environment, confinement systems that break agriculture's traditional ties to land and climate. This break makes confinement poultry production transferable to almost anywhere in the world, including developing countries. Thus, in developing countries the technology employed in confinement poultry is often more advanced than in other sectors of agriculture.

After an introductory overview, investment options for modern broiler production and processing in developing countries are outlined. A brief history of the spread of broiler production shows the transferability of the underlying technology. The paper concludes with data about the increasing dependency of many developing countries on imported feedstuffs to supply their expanding livestock industries.

Technology and Income Transform Agriculture

The transition of developing countries from food grain-based agriculture to include feed production and confinement raising of poultry and other livestock is a major long-term trend with rising incomes and urbanization. As incomes rise, people change their diet, usually including more livestock products. Not only does home consumption rise, with urbanization more products are marketed through fast-food shops, for which poultry products are well suited. Further, people are usually willing to pay more for fresh meat than frozen, imported meat. Thus, when personal incomes rise and the number of urban consumers is large enough to create a viable market, the technology for confinement poultry production readily transfers from developed countries for local production of chickens.

Gary Vocke is an economist with the Economic Research Service, U.S. Department of Agriculture.

Spread of Modern Poultry Production Requires Investment

Krostitz estimated in 1988 that in the higher-income countries of Latin America, North Africa, and the Far East, the proportion of chicken production in modern systems is between 70% and 90% of total output. In Middle Eastern OPEC countries, it is above 90%. Even in many of the lower-income countries in Asia and Africa, the proportion of the total output coming from modern operations is between one- and two-thirds. Krostitz also estimated that the investments for developing country poultry operations are about U.S. \$30 billion at today's prices.

The firm-level investments discussed here for a modern broiler operation in developing countries rely on a World Bank study of integrated broiler production. The minimum size considered by the World Bank produces 2.5 million broilers annually. Although 2.5 million is large, a Michigan State University study prepared at the same time assumed the minimum U.S. size to be 15 million broilers annually (Lasley).

Buildings for the various operations can be constructed using local materials and labor, so investment costs will vary by country. Most of the equipment, however, will likely be imported from international suppliers (see table 1).

In addition to the investments shown in table 1, large storage facilities will be needed where timely grain deliveries are uncertain. Further, if grains are purchased locally, drying equipment will be needed. If temperatures go over 95–100 degrees Fahrenheit frequently, the breeder and growout houses will need equipment to cool the air.

Another investment decision is the extent to automate. On average, one person using hand feeders and waterers can raise 7,000 broilers at a time. Automated feeding and watering increases the number of broilers per worker to 20,000, but raises equipment investment 280%. For comparison, fully automated U.S. facilities allow one person to raise 80,000 broilers at a time with only a half day of work. Automated

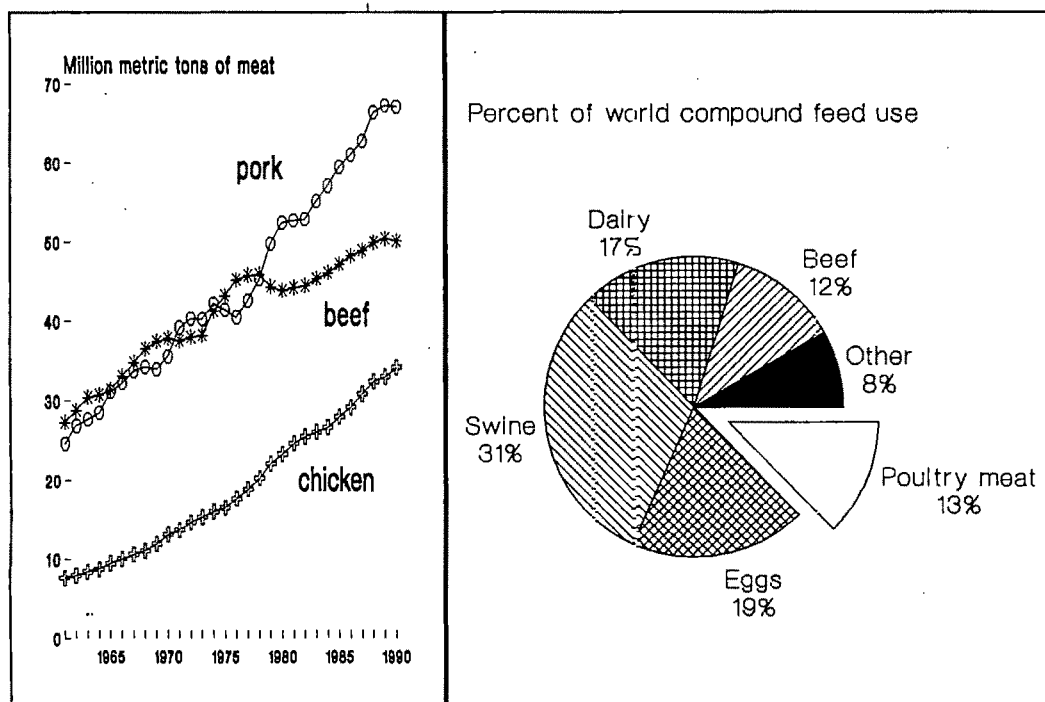


Figure 1. With present trends, poultry and swine will be increasing their share of world meat output and feed use (FAO, Vocke)

watering and feeding in breeder houses raises equipment investment 330%, but increases the number of layers per worker from 2,000 to only 2,500. There is less labor savings with breeder hens because most of the labor is for collecting the hatching eggs.

Facilities for feed milling, incubation of hatching eggs, and broiler processing require relatively large equipment investments and provide economies of size. For example, doubling output from 2.5 million broilers to 5.0 million

reduces equipment investment in these facilities by 25% per broiler marketed.

Technology Transfer Limits U.S. Exports

The technology for broiler production easily transfers, as has been shown again and again since U.S. scientific advances beginning in the 1940s allowed confinement production of poultry. These scientific advances gave U.S. farmers an early export advantage. Before U.S. agribusiness began exporting this broiler production technology in the 1960s to Western Europe and Japan, U.S. exports were more than 30% of world chicken meat trade. Though U.S. chicken meat exports are larger now than in the 1960s, the U.S. share has been cut nearly in half, to 17% (FAO).

The U.S. chicken meat exports to Europe began to decline with the technology transfer in the early 1960s as Dutch exports replaced U.S. exports to West Germany. Broiler production continued to increase in the European Community (EC), outpacing consumption. Now, led by France, the EC is a leading exporter of chicken

Table 1. Investment Breakdown for an Integrated System Producing 2.5 Million Broilers Annually

Facility	Total Investment by Facility	Investment for the Equipment
	(U.S. \$)	
Feed mill	300,000	150,000
Breeders	650,000	65,000
Hatchery	590,000	295,000
Growout	2,900,000	290,000
Processing	1,060,000	265,000
	5,500,000	1,065,000

Note: Adapted from the World Bank.

meat. France used government subsidies to establish broiler operations in Brittany because of that region's high unemployment (U.S. Department of Agriculture, FAS).

Japan's chicken production began rising with the transfer of broiler technology in the late 1950s. Consumption, however, outpaced local production because limited agricultural land and high population density restricts broiler production and processing. Thus, Japan remains a chicken meat importer from the United States and Thailand, and recently from Brazil.

Thailand began modernizing its broiler industry in the early 1970s. Although Thailand's production costs once were an estimated 30% higher than in the United States, imported technology and reorganizing producers from independent growers into integrated companies similar to those in the United States narrowed the U.S. advantage (U.S. Department of Agriculture). Using its low-cost labor to compete with U.S. exporters, most of Thailand's exports are deboned broilers to Japan.

Brazil also began exporting chicken meat in competition with the United States in the early 1970s, principally to the Middle East. Poultry production technology and organization in Brazil is similar to that of the United States. However, until 1988, Brazilian exporters benefitted from low-interest government loans and export subsidies (Ahmed). Although Brazil is now self-sufficient in poultry feed, many developing countries with rapidly increasing livestock production have had to import feedstuffs.

Chicken Production Often Leads to Imported Feed

Except in a few densely populated areas where shortage of land and pollution problems limit poultry production, the question of whether to import poultry meat or to build up a domestic poultry industry, even if feed must be imported, has generally been answered in favor of the building up of the poultry industry. The use of modern systems is higher in the Middle East than elsewhere in the third world because broiler production is relatively new in the region. Much of the region's poultry industry developed as incomes in the region rose following the high 1970s oil prices. The industry was created using the best technology, often transferred in large turnkey projects including the feed mill, breeder farms, hatchery, growout houses, and processing fa-

cility. One manufacturing company alone reported over 100 turnkey projects in the Middle East and North Africa from 1975 to 1985 ("Big D Achieve . . ."). A few countries provided incentives for this technology transfer. Saudi Arabia, for example, subsidized corn and soybean imports for broiler production.

Broiler production has also been growing rapidly in the developing countries of the Far East as incomes rose with exports of manufactured goods. In the Far East there was less importing of turnkey projects, more a steady upgrading of existing breeder and growout operations and shifting from live-bird markets (with on-the-spot slaughtering for the consumer) to factory processing.

The consequences of rising broiler production for feed trade vary. Many of the nations in these regions lack adequate cropland or suitable high-yielding crop varieties, requiring feedstuffs imports (fig. 2). For example, as the arid Middle East and North Africa reduced their dependency on chicken imports, dependency on feed imports increased. Elsewhere, even in countries in the Far East with abundant rainfall, feedstuff production is also often inadequate. For example, Malaysia's cropland and climate is not suitable to grow enough corn and soybeans, so the country must import part of its feed needs.

Without imports of feed ingredients, poultry producers in many of the countries in these regions, and elsewhere in the third world, would have to rely on high-cost locally grown feedstuffs, slowing the expansion of their broiler industries. Availability of low-cost feedstuffs is critically important. Feed will be 70%–80% of the total production cost of a broiler.

Poultry Trade Is Limited

One consequence of the easy and widespread transfer of confinement poultry production is that only 3%–4% of world poultry production is traded (fig. 3). This percentage is lower than the 15%–20% for some individual crops. Because crops are dependent upon both climate and suitable land, and not all countries are blessed with suitable resources, crop trade is greater.

Summary

Chicken meat consumption has grown very rapidly with rising incomes in the more prosperous

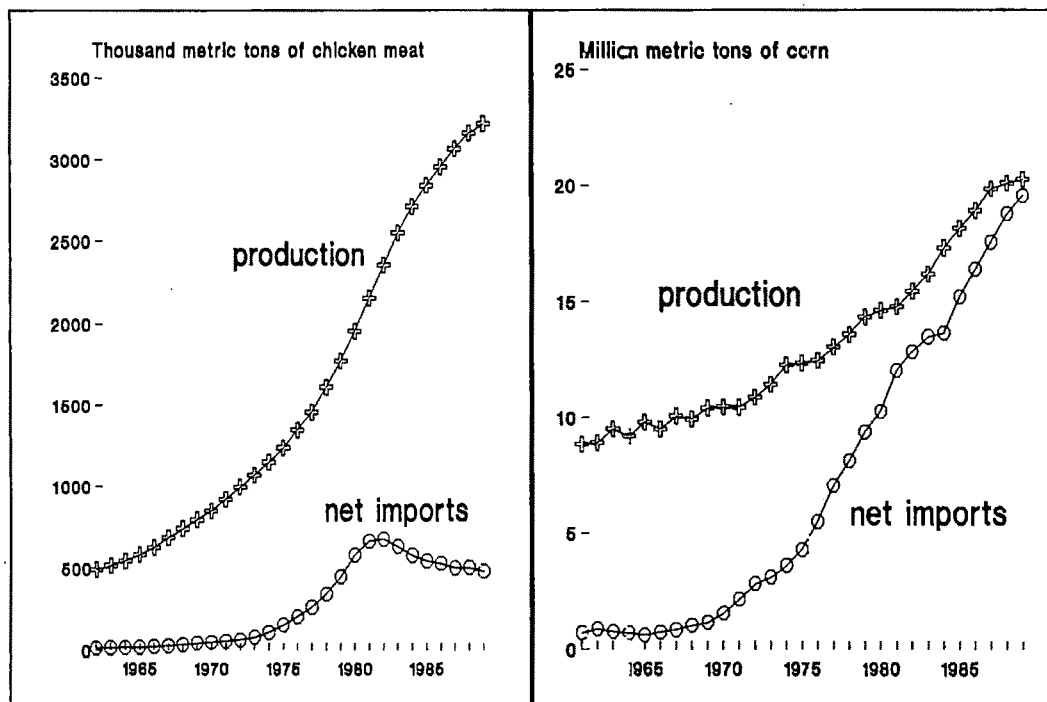


Figure 2. Production and trade of developing countries of the Middle East, North Africa, and Far East that import chicken meat (FAO)

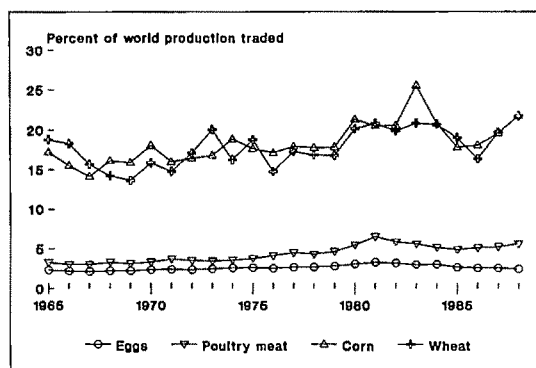


Figure 3. The percent of poultry production traded is less than crops (FAO)

developing countries, initially increasing imports of chicken meat. However, with rising consumption, incentives arise to import the technology to increase domestic production of chicken meat. This domestic production then substitutes for imports, slowing and sometimes reversing chicken meat trade. However, this transfer of technology often contributes to in-

creased imports of grains needed for modern livestock production.

References

- Ahmed, Hassan F. "Brazilian Poultry Industry: Situation and Outlook." *Dairy, Livestock, and Poultry: U.S. Trade and Prospects*. FDLP 1(1990):1-9.
- "Big D Achieve a Century of Projects With Acolid." *World Poultry* 10(1986):28-29.
- Food and Agriculture Organization of the United Nations (FAO). Production and trade data computer tapes. Rome, 1990.
- Krostitz, W. "The Poultry Boom in Developing Countries." *Zootecnica Int.* 5(1988):35-40.
- Lasley, Floyd, et al. *The U.S. Broiler Industry*. Washington DC: U.S. Department of Agriculture, Econ. Res. Ser., AER No. 591, 1938.
- U.S. Department of Agriculture, Foreign Agricultural Service. "Highlights of Major Poultry Competitors." *Livestock and Poultry: U.S. Trade and Prospects*. FDLP 2(1937):11-16.
- Vocke, Cary. "Egg Production: a Case of Successful Technological Transfer." *World Agriculture Situation and Outlook* 61(1990):8-9.
- World Bank. *Appraising Poultry Enterprises for Profitability: A Manual for Potential Investors*. Washington DC: World Bank Tech. Pap. No. 10, 1983.

International Capital Markets and Interest Groups

Terry Roe

Capital market integration has increased the scale and mobility of capital movements within and among nations to unprecedented levels. Integration is fairly recent. Frankel concludes from his review of the literature that barriers in the major world financial markets were sufficiently low that by 1989, financial markets could be viewed as being fully integrated.¹ The scale and mobility of capital flows among nations has numerous implications to currency values, interest rates, wealth, and, through current capital account linkages, to relative product and factor prices. Capital flows that have real economic effects also affect the functioning of policy instruments and how effective they are in remunerating an interest group at the expense, if any, of others.

This paper provides insights into the process and linkages by which the integration of capital markets has altered economic policy through interest group pressures on government. If economic policy is viewed as the outcome of interest groups seeking through the power of the state that which cannot be provided by the market alone, then capital market adjustments that affect the return to policy instruments can also affect the returns to lobbying resources and, hence, the balance of political influence among the various groups.² This type of explanation for economic policy has been forwarded under the rubric of models of "rent seeking" (Krueger 1974).³

Terry Roe is a professor, Department of Agricultural and Applied Economics, University of Minnesota.

¹ However, evidence suggests that increased volatility in foreign exchange markets has tended to mask comparative advantage in the case of some countries and limited the degree of integration for financial instruments of longer duration (Artis and Bayoumi).

² Lobbying is used as a surrogate for voting, donations to political action committees, and other means of influencing a government's choice over policy instruments. The key point is that resources are required to influence choice and that interest groups behave according to what they perceive to be in their own self-interest.

³ Other labels are "directly unproductive profit-seeking activity" (Bhagwati) or more generally "political economy" (Colander). Because numerous guides to this literature can be found easily, we provide no review here.

The paper is organized as follows.⁴ Selected historical evidence linking capital markets to pressures for legislation in the United States and other countries is briefly reviewed. In some cases, capital market adjustments have eased pressures for protection, while, in other cases, adjustments either forced policy change because of a liquidity constraint or made unprofitable the lobbying for sustaining policy in face of these adjustments. The discussion in the next section provides deeper insight into capital market-lobbying linkages. Discussion focuses on the results from a simple model of rent seeking, where government intervenes in markets and provides sector-specific public goods.

Interest Groups and Capital Market Adjustments: An Overview

The linkage between the overvalued dollar during the Bretton Woods period and the development of agricultural policy was discussed by Schuh in his seminal paper on exchange rates and agriculture. Under Bretton Woods, the overvalued dollar served implicitly to tax producers of agricultural tradables. Farm organizations grew into potent lobbying groups during this period. Their success in obtaining legislation to support farm incomes is evident. The influence of capital markets on pressures for farm legislation during the pre- and post-Bretton Woods periods has also been linked. In a study of the origins of legislation to protect farm incomes, Goldstein concludes that the economic environment of the 1920s helped motivate the formation of agricultural lobbying groups when it became generally apparent to rural groups that the post-World War I economy was discriminating against agriculture relative to other sectors of the economy.

⁴ See the background paper to this article by the same title for a discussion of capital market integration and current-capital account linkages in an open capital market economy.

In a review of the post-World War I economy, Orden suggests that these pressures would likely have been even stronger had capital markets not alleviated the otherwise larger appreciation in the U.S. dollar in response to reparation of European debt. Following World War I, capital markets provided reverse capital flows from the United States to Europe. If these reverse flows had not occurred, Europe would have maintained an even larger trade surplus. The U.S. dollar would have appreciated even further to restrict U.S. exports thus placing an even greater implicit tax on U.S. agricultural exports.

The effectiveness of capital markets to meet U.S. demand for loanable funds in the late 1980s has helped the United States sustain its budget deficit, albeit with the burden that debt service payments eventually imply. The burden of servicing increasing debt can be accomplished by a real depreciation of the currency, and/or a recession that reduces imports. If investors perceive an increase in exchange rate risk, real interest rates could rise. Perhaps equally alarming, the outstanding debt decreases the flexibility to respond to a recession because lowering taxes or increasing expenditures contributes to debt and threatens the willingness of investors, at the margin, to hold or accumulate more of the country's assets. Federal Reserve policy is also constrained because expansionist monetary policy will tend to depreciate the currency and raise inflation fears, while lower interest rates would contribute further to a decline in foreign capital inflows.

Were special interests cognizant of these factors while lobbying for policies that prevented adjustments in the budget deficit? Capital inflows possibly lessened pressures for additional farm subsidies. Still, the high real interest rates and the value of the dollar contributed to a decline in land values and a decline in the U.S. share of world agricultural trade. Legislation in response to these developments included the Export Enhancement Program. Ironically, as the 1990 farm bill suggests, the implication of growth in the U.S. budget deficit for action under Gramm-Rudman⁵ and higher taxes appear to have altered the political balance to favor a decrease in income support to agriculture, in spite of the apparent unwillingness of the European Community (EC) to liberalize agricultural trade sig-

nificantly.⁶ Hence, it seems reasonable to infer that world capital markets allowed the United States to delay adjustment to its growing budget deficit. Then, given the magnitude of the deficit, the rather modest adjustment that occurred led to a trade-off among special interests that induced the United States to partially liberalize agricultural trade.

Empirical support for the hypothesis that a strong dollar has precipitated protectionist legislation in the nonagricultural sector of the U.S. economy is judged by Leamer as less than conclusive. Perhaps the strongest evidence linking protectionist legislation and the appreciation of the U.S. dollar is provided by Grossman's study of the U.S. steel industry. Grossman finds that the steel industry's request for trade protection coincided with an appreciation of the dollar. The appreciating dollar and the corresponding loss of competitiveness was found to account for over one-half of the approximately 16,000 jobs lost in the industry.

The linkage between capital markets and interest groups is perhaps more evident in developing countries that have pursued inward-oriented economic policies (Balassa 1986). The World Bank-sponsored study of the political economy of agricultural price policies in eighteen countries found that an overvalued currency tended to account for larger indirect taxes on producers of agricultural exports than taxes imposed by interventions in the sector alone.⁷ Countries with higher indirect taxes were associated with those that exercised control over capital markets and pursued various types of exchange-rate-control regimes. Korea and Malaysia, whose capital markets were relatively open, tended to discriminate far less against agriculture than did the other countries (World Bank).

Controls on domestic capital markets are typically coincident with import-substitution-industrialization policies. Interventions in domestic capital markets have given rise to negative real interest rates in part because nominal rates re-

⁶ Johnson, Mahe, and Roe found support for the hypothesis that U.S.-E.C. agricultural policies pursued in 1986 could be characterized as Nash equilibria. A movement from observed policies, while saving resources, made both "politically" worse off as measured by an empirically estimated policy goals function. Preliminary results for the 1988 base year suggest the same conclusion. A departure from this equilibrium is possible if interest group pressures increase the political importance of the budget relative to income transfers to farmers, as the new legislation implies.

⁷ See Kueger, Schiff, and Valdes for a synthesis of some of the results of these studies. Individual country results appear in the World Bank's "Comparative Country Study Series on the Political Economy of Agricultural Pricing Policies."

⁵ See the 1990 Budget Reconciliation Act for the modifications to the 1990 farm bill.

main fixed during periods of high inflation. At various times during the 1970s, negative real interest rates were particularly large in Brazil, Ghana, Jamaica, Nigeria, Peru, and Turkey. The experiences documented by Balassa (1984) for Brazil and Turkey and by Corbor, deMelo and Tybout for the Latin American countries of the Southern Cone support the view that the effects of below-equilibrium real interest rates tended to lower the efficiency of investment by discriminating among capital users, to favor the allocation of rationed capital to capital-intensive techniques, to discourage domestic savings, and to encourage capital outflows.

Reform in many countries included in the IBRD study was forestalled when rising world prices for exportables increased liquidity. The constituencies supporting trade regimes became entrenched. When adjustment was needed, negotiations between various interest groups either proceeded slowly or did not occur in time to alleviate a liquidity crises. The result for many countries was an externally devised liberalization program that likely would not have come about from a realignment of interest groups alone.

The balance of political power to influence policy in most of these countries appeared to be urban based. Urban-based coalitions tended to support the use of the exchange rate, the trade and payments regime, and schemes to allocate credit at low nominal and negative real interest rates, all of which tended to discriminate against agriculture.⁸ If domestic capital markets had been open to world markets, exchange rate adjustments would have at least partially undone the objectives of sector interventions because currency depreciation would have raised the real price of food to consumers and adjustment in relative prices would have decreased the discrimination against agriculture relative to the nonagricultural sector.

There are both economic and political equilibria. An extraordinarily large economic adjustment that may lessen market distortions may so threaten or realign the balance of political influence that investors (asset holders) perceive increased risks to political stability and to asset returns. This tendency, coupled with the increased investment opportunities created by the ever more integrated world capital markets, likely increases the incidence of capital flight.⁹ Hence,

the increased capacity of world capital markets to more cheaply and quickly supply capital to perceived investment opportunities can increase both the cost of realigning economic policy and the cost of policies that induce domestic market distortions.

A More In-Depth, Narrow View of Special Interests and Capital Markets

To provide some insight into the economic forces that influence political equilibria, it is useful to structure the discussion around a model of rent seeking. Even though the model excludes capital markets and imported intermediate capital goods, some implications of the current account–capital account linkages to the welfare of households can be discerned. The model resembles that of Young and Magee, except for purposes here, we focus on the individual household and treat the political process as a black box.¹⁰ Consider a small two-sector economy that produces and consumes an export and an import good using only labor L_i , a sector-specific factor x_i , and a sector-specific public good G_i , where the sector index is $i = r$ (rural), u (urban). The government's policy instruments are the relative price p of the export good and the production of public goods using the technology $G_i(l_i^g)$, where l_i^g denotes government hired labor. The price of the import good is *numéraire*.

Let the typical household's "conditional" indirect utility function be denoted as

$$(1) \quad V_i[p, \pi_i[p, w, c_i, G_i] + w(\bar{L}_i - l_i) + c_i \bar{x}_i + \gamma_i T[G, p, p^w]],$$

where $\pi_i(\cdot)$ is the indirect profit function; w and c_i are wages and the price of a factor that is only traded within the sector (e.g., land), labor and sector-specific endowments are \bar{L}_i and \bar{x}_i , respectively; γ_i is the proportion of the total tax bill T paid by the i th household; and p^w is relative world market price.¹¹ Taxes depend on the government's expenditures on public goods G_i and the fiscal costs associated with interventions that cause a departure between domestic and

sued. Uncertainty as to whether a new regime can be sustained leads asset holders to slow the repatriation of flight capital from markets outside the country.

¹⁰ The framework is adapted from Roe and Graham-Tomasi. They provide a more detailed specification of the political process and show the derivation of the type of results reported here.

¹¹ The tax equation prevents the government from obtaining "free" resources from the rest of the world.

⁸ See Greene and Roe for the case of the Dominican Republic.

⁹ Dornbusch's study of the Mexican experience with stabilization policy strongly suggests a political motivation for the policies pur-

world prices. The function is referred to as conditional since the lobby level l_i is the remaining variable to be chosen. Presumably, lobbying influences the government's choices, G and p .

Before the household can optimally choose l_i , it is reasonable to assume that it holds beliefs about the response of the political process to its lobbying efforts. That is, the household must hold beliefs about the government's policy decision rules. Let these rules be denoted as

$$(2) \quad p = p(l_r, l_u; Z),$$

$$(3) \quad G_i = G_i(l_r, l_u; Z)$$

where Z denotes a vector of exogenous variables, such as world price p^w , and resource \bar{L}_i , \bar{x}_i , institutional and political endowments, and so on. In general equilibrium, factor prices w , c_i are endogenous, they depend upon p , p^w , G_i , the other mentioned exogenous variables.

The necessary conditions for l_i to maximize (1.0) appear in table 1. Essentially, the condition indicates that l_i is chosen to equate marginal returns with marginal costs. To illustrate, consider term 1. If the rural household produces in excess of own consumption q_{rr} , ($y_r - q_{rr}$) positive, and its lobbying efforts result in an increase in the price of y , relative to the urban good, $\partial p / \partial l_r$ positive, then the household realizes a gain from lobbying. Thus, the more specialized is the household, y_r large relative to q_{rr} , the more willing it is to allocate resources to

influence price policy, all else constant. The availability of a cost-reducing technology G_r , term 3 positive, also increases the household's willingness to influence policy.

The trade-off between rural and urban households can be seen by considering the equivalent of term one for the urban household. This term can be shown to be $-q_{ru} \partial p / \partial l_u$, where the urban household lobbies for a decrease in the price of food, i.e., $\partial p / \partial l_u < 0$. Hence, all else constant, rural households have an incentive to lobby for an increase in the price of the rural good while urban households have an incentive to lobby for a price decrease. Depending on the sign of $\partial^2 p / \partial l_i \partial l_j$, they may also be willing to countervail the lobbying efforts of the other. A prisoner's dilemma is one possible outcome.

If $(\bar{L}_r - L_r - l_r)$ positive, rural household labor is employed in the urban sector. In the Ricardo-Viner type of model specified here, a component of the derivative dw/dl_r is $\partial w / \partial p (\partial p / \partial l_r)$ where $\partial w / \partial p$ is positive. Then, the household has an additional incentive to lobby for a price increase since real wages rise. If another freely mobile input were specified, then the sign of $\partial w / \partial p$ depends on the Stolper-Samuelson condition.

A number of points can now be made. First, even though capital markets do not appear in the model, it can be seen that current and capital account linkages that affect relative prices of commodities or factors can induce households

Table 1. Necessary Conditions for Determining the Household's Willingness to Pay to Influence Economic Policy

Term 1, ^a Commodity Market		Term 2, Labor Market		Term 3, Public Good
$(y_r - q_{rr})p_{lr}$	+	$(\bar{L}_r - L_r - l_r) w_{lr}$	+	$\pi_{r,G_r} G_r, l_r^g, l_r$
The product of household production y , less consumption q_{rr} , i.e., market surplus, and the change in price p from additional lobbying l_r .		The product of the household's net labor positions and the indirect effect of the change in wages w from additional lobbying l_r .		The product of the shadow price π_{r,G_r} of G_r , its marginal product G_r , l_r^g , and the change in the amount of government labor employed, l_r^g, l_r .
Term 4, Sector Specific Factor Market		Term 5, Taxes		Term 6, Opportunity Cost of Lobbying Res.
$(\bar{x}_r - x_r)c_{r,l_r}$	+	γT_{lr}	-	$w = 0$
The product of the household's net employment position of the sector specific factor and change in its price from additional lobbying		The change in the household's tax bill from additional lobbying.		Wage, the opportunity cost of lobbying.

Note: See Roe and Graham-Tomasi for the details of this derivation.

^aNotation p_{lr} and π_{r,G_r} , for example, denotes the partial derivatives $\partial p / \partial l_r$ and $\partial \pi_r / \partial G_r$, respectively.

to alter their lobbying levels depending on their preferences, distribution of endowments, and the competitiveness of one sector relative to the other. Effectively, they act as though they choose l_i to re-equilibrate the first-order marginal return–marginal cost condition.

Consider, for example, the sector-specific factors such as land, plant and equipment, term 4. Policies that have an adverse impact on the value, c_i , of sector-specific endowments can also have an adverse impact on the sector's capacity to invest when \bar{x}_i meets collateral requirements. Policies that discriminate against agriculture can implicitly decrease its capacity to obtain credit relative to protected sectors. Further, just as the value of protection gets built into the value c_i of sector-specific assets, so too do the effects on c_i attributable to the performance of capital markets. Rent is $\bar{x}_i c_i$, which determines the value of this asset and, hence, wealth. If \bar{x}_i were plant and equipment of a sector protected using capital market controls of the nature mentioned, then liberalization can imply a decline in wealth of large magnitudes. If the household correctly perceives this possibility, then it may be willing to increase significantly its lobbying commitment to preclude such an eventuality. The distribution of endowments \bar{x}_i within the sector will also affect households' willingness to lobby differentially, depending on whether they are net suppliers or demanders of the sector-specific asset, $(\bar{x}_i - x_i)$ positive or negative, respectively. This has implications for sustaining a coalition.

Second, because the capacity, particularly of developing countries, to provide public goods (infrastructure, communications, education, public services) often depends on their ability to obtain loanable funds from abroad, domestic policy that gives rise to an efficient allocation of the country's resources is likely to increase its capacity to access international capital markets. The provision of public goods increases the productivity of private sector resources and the willingness of households to invest in quasi-fixed factors of production.¹² Countries following inward-oriented economic policies often incur substantial fiscal deficits, so that the marginal social cost of an incremental investment in public goods is extraordinarily high, all else constant. Further, public good allocation tends to have an urban bias (Braverman and Kanbur).

In this case, the framework posited in table 1 suggests that the household is concerned with

the optimal allocation of its lobbying resources so that, at the margin, it does not necessarily prefer the provision of public goods over policies that distort prices. In this case, the short-run marginal cost–marginal return to the household's lobbying resources might favor lobbying for price distortions as opposed to lobbying for the provision of public goods, particularly when households do not have access to world capital markets so that the provision of public good must be financed from domestic savings.

Three corollaries follow. First, to the extent that world capital markets have increased the incentive for capital to flee inward-oriented economies, then the cost of pursuing an inward orientation is increased further because, all else constant, the difficulty of financing the provision of public goods is increased as a result of the decline in domestic loanable funds and the adverse impact on imbalances in external accounts that capital flight implies. Declines in the provision of public goods clearly has dire implications for economic growth.

A second corollary of this argument is that when capital markets work imperfectly, as they almost surely did when many countries accumulated debt levels that later led to liquidity problems, households may view the tax component T in table 1 to be relatively unaffected by lobbying efforts, at least in the short run. Households may then continue to lobby for policies which, for some sustainable period, allow expenditures to exceed income. Effectively, governments, responding to the influence of special interests, may then pursue policies as though a financial constraint was not binding.¹³ The eventual burden of servicing external debt, as appears to have happened for many countries, has forced them to liberalize trade and other policies that contributed to domestic distortions. But, capital inflows to these countries have not returned (IMF). They now find themselves in a position where policy liberalization has occurred but implicit taxes on the consumption of traded goods is higher than it would likely be if capital inflows returned, i.e., their currencies are undervalued. Moreover, the cost of loanable funds to allocate to the provision of public goods is also higher in this environment.¹⁴

A third corollary is that the biased provision

¹³ See Mohtadi and Roe for the implications to economic growth when households ignore the tax implications of their lobbying activity.

¹⁴ An interesting conjecture is whether this line of reasoning also applies to the problems the U.S. will face in maintaining investments in public goods while servicing debt.

¹² See Binswanger for estimates of the effects of public good supply on the aggregate elasticities of agricultural supply.

of public goods can increase the willingness of the sector receiving more than its share to lobby for its differential advantage. The provision of the public good increases the willingness to invest in quasi-fixed factors of production, which, in turn, increases the first component of term 1; and, depending on the economies of scale of private and public goods in production, it can also increase the shadow price of the public good in term 3, table 1. Thus, the biased provision of public goods can increase the capacity of one sector to lobby for policy that further discriminates against the other sector.

The third point relates to the diversification of household assets and the functioning of domestic markets. This point also relates to Mancur Olson's suggestion that broad-based coalitions tend to take into account the adverse macroeconomic effects of their lobbying efforts so that the adverse effects of the differential advantage they seek tends to be less than those of narrow-based coalitions. The simple model in table 1 shows that, if households hold shares in both sector's sector-specific assets, then the household would have less incentive to seek its differential advantage because lobbying that benefited one sector at some cost of the other would be borne by both households. Households would have a greater incentive to cooperate, i.e., to lobby in ways that benefit both.

Capital markets permit households to diversify assets. In the presence of security markets, households can hold shares in the assets of other sectors and other countries, both directly and through retirement funds managed by capital market firms. Through diversification, these markets help households insure against policy (i.e., the influence of special interests) that adversely affects a sector or country. Hence, it can be conjectured that, all else constant, diversification increases the household's perceived cost of seeking its differential advantage. In terms of the simple model, the household may find it optimal to lobby for the provision of public goods as opposed to price distortions.¹⁵

The fourth point is closely related to the above. It deals with the presence of a lobbying equilibrium. More specifically, the seeking of differential advantage can be characterized as a non-cooperative game. A Nash solution can be

defined as

$$(4) \quad \partial V_r / \partial l_{r|l_r, l_r^*, z} = 0, \partial V_u / \partial l_{u|l_u, l_u^*, z} = 0,$$

where l^* are values that maximize (1) and $l_i(\cdot)$ are the rules obtained from a solution to the condition expressed in table 1. Condition (4) characterizes a stable political process; it allows (2) and (3) to be stated as functions of exogenous variables only. Of course, there may not be a Nash solution, and if there is, it need not be unique, nor is it the case that Nash behavior is most natural here.

Institutions are important because they are a component of the political technology available to respective households, the black box behind this model. This technology is important in determining the nature of a political equilibrium, i.e., the lobbying levels for which neither household is willing to change. Note, too, the knowledge and information requirements of the household, the prerequisites for informed, rational choice. A competitive market model with no government is informationally efficient because households need to respond only to the information provided by price signals. However, in the case considered here, the household's decision-making problem is more complex because of knowledge of the political "technology," equations (2) and (3), the effects of lobbying on market-determined variables, w , c , and c_s , taxes T , and the countervailing lobbying of adversely affected households. As noted by others, e.g., Krueger (1990, p. 21) "when the costs of a policy are obscure, special interests in the private sector and government have a greater opportunity to use those policies for their own advantage without incurring the disapprobrium of voters and other politicians."

Concluding Comments

Integration of capital markets has increased the scale and mobility of capital movements within and among nations. This paper has attempted to provide some insights into the process and linkages by which integration has influenced interest groups and hence economic policy. It is suggested that capital market adjustments and pressures for farm legislation can be traced back to post-World War I, although similar evidence for the industrial sector seems less well documented. It is conjectured that the balance of political influence delayed adjustments to the emerging U.S. budget deficit in part because world capital markets supplied the U.S. demand

¹⁵ If sector-specific factors x_i can be freely traded and labor is permitted to migrate freely between countries at low cost, much as in the case of countries within the European community, then models of the type posited here indicate that returns to lobbying in a particular country that tax some resources relative to others tend to decline.

for loanable funds at low cost. When implications of the deficit became more apparent to a broader range of interest groups, then political balance favored a decrease in support for agriculture in spite of the EC's reluctance to liberalize trade to anywhere near the same degree.

The use of capital market controls to discriminate against agriculture is most apparent in many of the world's developing and newly industrialized economies. Had these economies been open to the world's capital markets, then domestic capital market adjustments would have at least partially undone the objectives of sectoral interventions, hence the motivation for special interests to lobby for capital market controls.

More in-depth insights, and insights supportive of the previous discussion, are provided using a simple model of rent seeking. Key points are that capital markets affect the value of sector-specific resources and, hence, the value of collateral and wealth as determined by these assets. Collateral is associated with investment and the productivity of variable resources. Productivity growth sets in motion a chain of events including higher returns to resources and the specialization and division of labor. This chain of events also alters a household's willingness to lobby. If policy discriminates against the sector, then this chain of events can induce households to lobby for the removal of distortions. The converse is also possible. Because capital market adjustments affect the wealth embodied in sector endowments, households may be willing to expend lobbying resources to resist those capital market adjustments that adversely impact on wealth, e.g., the protection overvalued exchange rates and negative interest rates afford the industrial sector of many inward-oriented economies.

Countries following inward-oriented policies often incur substantial fiscal deficits so that the marginal cost of an incremental investment in public goods is extraordinarily high. In this case, the marginal return-cost calculus of the household may favor lobbying for interventions that distort markets rather than for interventions that increase the provision of public goods. However, to the extent that capital markets have increased the incentive for capital to leave inward-oriented economies, the cost of pursuing an inward orientation is increased. Effectively, capital flight increases the likelihood of a binding liquidity constraint making the maintenance of an inward-oriented policy less feasible regardless of the willingness to lobby for their support, unless of course, autarky is preferable.

A corollary is that capital markets may make fiscal and trade imbalances appear artificially cheap, thus inducing some households to ignore the tax implications of their lobbying efforts.

Capital markets permit the internalizing of lobbying externalities when households rely on these markets to diversify their portfolios. Then, less incentive exists to seek a differential advantage since lobbying that benefited one sector at some cost to the other is borne by the household through its portfolio linkage. This effect, too, may be a two-edged sword if households are not fully informed. Capital inflows to the United States during the 1980s appear to have contributed to rising security prices. Did rising security prices lead to additional support for supply-side policies?

The rather complex and obscure dimensions of capital market effects on an economy surely increase the complexity for the household to lobby for policies that truly increase its welfare. When the costs of policy are obscure, the lobbying costs incurred by special interests tend to be lower because of the lack of countervailing efforts by others. Political technology, ideology, knowledge, and information are particularly important to effect rational choice and a balance of competition in the "market place" for public choice. Institutions, a public good, seem to be particularly important in this regard.

The evolution of world capital markets should, over time, decrease the incentives for interest groups to induce governments to pursue policies that discriminate in their favor. This is another way of suggesting that evolution will increase the difficulty of pursuing policies that cause an inefficient allocation of resources and that externalize costs to other countries. This evolution also increases the importance of international institutions and agencies to perform activities that permit markets to function efficiently, such as enforcing voluntarily negotiated contracts, provision of information, and regulating sources of instability. Finally, efforts to liberalize an economy may focus first on capital markets because these markets impact on the entire economy and because of the second-best problems associated with sectoral liberalizations.

References

- Artis, Michael, and Tamim Bayoumi. "Saving, Investment, Financial Integration, and the Balance of Payments." *Staff Studies for the World Economic Outlook*. International Monetary Fund Staff Stud., Washington DC, Sep. 1990.

- Balassa, Bela. "Adjusting to External Shocks: The Newly Industrialized Developing Countries in 1974-76 and 1979-81." Washington DC: World Bank Disc. Pap. No. DRD89, May 1984.
- . "Policy Responses to Exogenous Shocks in Developing Countries." *Amer. Econ. Rev.* 76(1986):244-48.
- Bhagwati, N. Jagdish. "Directly Unproductive, Profit-Seeking (DUP) Activities." *J. Polit. Econ.* 90 (1982):988-1002.
- Binswanger, Hans. "The Policy Response of Agriculture." *Proceedings of the World Bank Annual Conference on Development Economics 1989*. Washington DC: World Bank, 1990.
- Braverman, Avishay, and Ravi Kanbur. "Urban Bias and the Political Economy of Agricultural Reform." *World Develop.* 15(1987):1179-87.
- Colander, David C., ed. *Neoclassical Political Economy: The Analysis of Rent Seeking and DUP Activities*. Cambridge MA: Ballinger Publishing Co., 1984.
- Corbo, Vittorio, Jaime de Melo, and James Tybout. "What Went Wrong with the Recent Reforms in the Southern Cone." *Econ. Develop. and Cultur. Change* 34 (1986):607-40.
- Dornbusch, Rudiger. "The New Classical Macroeconomics and Stabilization Policy." *Amer. Econ. Rev.* 80(1990):143-47.
- Frankel, Jeffrey. "Quantifying International Capital Mobility in the 1980s." *Saving*, ed. Doug Bertheim and John Shoren. Chicago: University of Chicago Press, 1989.
- Goldstein, Judith. "Ideas, Institutions, and American Trade Policy." *Int. Org.* 42(1988):180-217.
- Greene, Duty, and Terry L. Roe. "Trade, Exchange Rate, and Agricultural Pricing Policies in the Dominican Republic." *The Political Economy of Agricultural Pricing Policy*. Washington DC: World Bank Comparative Studies, 1988.
- Grossman, Gene M. "Imports as a Cause of Injury: The Case of the U.S. Steel Industry." *J. Int. Econ.* 20(1986):201-23.
- International Monetary Fund. *International Capital Markets: Developments and Perspectives*. Washington DC, April 1990.
- Johnson, Martin, Louis Mahe, and Terry Roe. *Politically Acceptable Trade Compromises between the EC and the US: A Game Theory Approach*. International Agricultural Trade Research Consortium Work. Pap. No. 90-5, Dep. Agr. and Appl. Econ, University of Minnesota, 1990.
- Krueger, Anne O. "Government Failures in Development." *J. Econ. Perspectives* 4(1990):9-24.
- . "The Political Economy of the Rent-Seeking Society." *Amer. Econ. Rev.* 64(1974):291-303.
- Krueger, Anne O., Maurice Schiff, and Alberto Valdes. "Agricultural Incentives in Developing Countries: Measuring the Effect of Sectoral and Economywide Policies." *World Bank Econ. Rev.* 2(1988):255-71.
- Leamer, Edward E. "Comment on Endogenous Protection in the U.S., 1900-1984." *US Trade Policies in a Changing World*, ed. Robert M. Stern, pp. 196-200. Cambridge MA: MIT Press, 1987.
- Mohtadi, Hamid, and Terry Roe. *Political Economy of Endogenous Growth*. Paper presented at Amer. Econ. Assoc. meeting, Washington DC, Dec. 1990.
- Orden, David. "International Capital Markets and Structural Adjustment in U.S. Agriculture." *Amer. J. Agr. Econ.* 72(1990):745-54.
- Roe, Terry, and Theodore Graham-Tomasi. *Competition among Rent Seeking Groups in General Equilibrium*. Dep. Econ. and Dep. Agr. and Appl. Econ. Develop. Ctr. Bull. No. 90-2, University of Minnesota, Sep. 1990.
- Schuh, G. Edward. "The Exchange Rate and U.S. Agriculture." *Amer. J. Agr. Econ.* 56(1974):1-13.
- World Bank. *The Political Economy of Agricultural Pricing Policies*. Comparative studies of various countries, Washington DC, 1989.
- Young, Leslie, and Stephen P. Magee. "Endogenous Protection, Factor Returns and Resource Allocation." *Rev. Econ. and Statist.* 53(1986):401-19.

International Capital Markets and Development Funds for Agriculture: Discussion

Ralph W. Cummings, Jr.

The contributed papers do an admirable job of describing the new environment which has emerged over the past quarter of a century in which developing countries can supplement their own resources by tapping the international pool of savings now available from a huge, well-integrated capital market.

Roe's paper provides a framework to analyze the conditions—checks and limitations—imposed on economic decision makers by the international capital market. He argues that the evolution of world capital markets should, over time, decrease incentives for interest groups to induce governments to pursue policies that discriminate in their favor. This evolution increases the importance of development agencies to support activities that permit markets to function efficiently. Economic equilibria are interdependent with political equilibria; economic reforms must be sensitive to political stability.

Vocke describes the transfer of intensive poultry production systems. These operate very much the same whether in the United States or Thailand. This is a clear case where international capital markets should be tapped directly.

The paper by Lele synthesizes an ambitious systematic examination, using detailed case studies rather than polemics or econometric manipulations, of the role of foreign assistance in development. It points out some of the shortfalls of foreign assistance, among which instability, unpredictability, and biasing of investment decisions to availabilities of funds from donors are prominent. It underlines how the debt crisis has severely curtailed the previously increasingly important net flows of private capital to developing countries.

As noted in Lele's paper, all is not well with

official development assistance. Private capital flows will have to start growing again if the demand from developing countries is to be satisfied. These papers have several implications for donors.

One implication is the need to get back to basics; that is, to look first to analyze whether countries have progressive agricultural sectors. One of the failures of our profession is that we have been unable to convince leaders in developing countries and in our own development agencies of the fundamental lesson that a progressive agricultural sector is the driving force for economic development.

A second implication is the need to improve stability in addressing comparative advantage in providing assistance to accelerate agricultural development. If countries have sound economic environments, they should be able to borrow for infrastructure, either from the international capital market (Roe notes that exchange rate uncertainties cause international capital markets to be reluctant to finance longer maturities) or from the multilateral development banks. If a donor country has a competitive advantage in supplying infrastructure, that should be expressed through competitive bidding.

Donors have a role in promoting a sound economic structure. Lele concludes that even small amounts of aid—given in the context of a sound policy environment for the development of physical, human, and institutional infrastructure—are far more effective than large amounts given without attention to the overall development policy. Roe argues that domestic policy that facilitates efficient allocation of the country's resources is likely to increase its capacity to access international capital markets.

Donors can divide responsibilities to promote economic reforms. The multilateral banks and other donors can provide loans to implement structural adjustments and focus on "big policy," that is, macropolicy. The comparative ad-

Ralph W. Cummings, Jr., is an agricultural economist with the Agency for International Development.

The views expressed in this discussion are the author's and do not necessarily reflect those of AID.

vantage of a bilateral donor like the Agency for International Development (AID) that gives a limited amount of assistance the old-fashioned way—as grants or as loans whose terms approximate grants—is in supporting analysis to ask questions and to provide answers about what to change and how to change it, with particular attention to “small policy,” for example, specific challenges such as facilitating technological change, private sector growth, and environmental responsibility (ensuring benefits for future as well as present generations). Roe points out that capital markets permit households to diversify assets; this should be encouraged because the differential advantage that broad-based coalitions seek tends to be less than that sought by narrow-based coalitions. Broad-based coalitions are more likely to support government expenditures for public goods. He also cites Krueger: “When the costs of a policy are obscure, special interests in the private sector and government have a greater opportunity to use those policies for this own advantage without incurring the disapprobrium of voters and politicians.” Improving the legal environment—commercial law, regulations, grades and standards, market information, property rights—may be a key way to stimulate broad-based private sector activity.

The systematic examination of the role of foreign assistance in economic development demonstrates that the United States has a clear comparative advantage in technology development,

particularly in human resource development, research, and training. Lele’s paper provides strong evidence of the important contribution to Asia’s economic successes, and, conversely, the absence of the contribution partially accounting for Africa’s failures. There are a number of studies which demonstrate over a range of countries and commodities that returns to research, if done well, are two to three times other agricultural investments. Advances in U.S. agriculture, making use of university, the U.S. Department of Agriculture, and private sector experience, provide an extraordinary example of the successful development and application of science and technology. And the United States can benefit from it. Generally, developing countries are reluctant to borrow from either private capital markets or multilateral development banks for human resource development. The benefits result after a long gestation period, but the costs come early. The private sector cannot easily recapture the costs. And the effort must be done right and sustained for at least ten to twenty years. Science and technology is what developing countries most need and what U.S. assistance does best.

Official development assistance is not expected to increase significantly. The newly emerging international capital market introduces a potentially very positive dimension to economic development prospects. All donors need to reconsider how they can best facilitate the contribution.

Books Reviewed

Anderson, Kym. *Changing Comparative Advantages in China: Effects on Food, Feed, and Fibre Markets*. Paris: OECD, 1990, 118 pp., price unknown.

This book was motivated by the enigmatic nature of the changing role of agriculture in China's modern economy. Much of Anderson's work over the course of his career has focused on the rise and fall of agriculture in small Asian countries, and this new book analyzes the big picture in Asia; namely, what does the future hold for Chinese agriculture? He argues that China is losing its comparative advantage in agriculture; but, in return, its comparative advantage in lightly manufactured products is increasing. These are long-term trends the author has identified, and he fully expects them to continue. He is not surprised by what he finds because he argues that China is following basically the same development path taken by other East Asian economies (e.g., Japan, Korea, and Taiwan).

Chapter 1 is an executive summary of the book. Chapter 2 provides the theoretical framework for the book's thesis. It offers an explanation of (a) "why the agricultural sector's shares of GDP and employment tend to decline as an economy grows" (p. 17) and (b) "why densely populated countries also tend to lose their agricultural comparative advantage as they develop" (p. 17). The explanation of (a) is much more complete than for (b). Anderson argues that in a developing economy, agriculture's share of national product normally falls because, first, productivity growth is faster in agriculture than manufacturing; second, demand for food grows slowly; third, growth in the rest of the world lowers the terms of trade for agricultural exporters; and, fourth, agriculture's share of tradables' production declines. He then argues that densely populated countries eventually lose their comparative advantage in agriculture because development increases the availability of capital and the ratio of capital to natural resources rises. Comparative advantages shift with changes in relative factor endowments. Drawing on empirical evidence from economies in East Asia that are more advanced than China, Anderson concludes that "these densely populated, resource-poor countries lost their comparative advantage in primary products at a relatively early stage of their economic development, and enjoyed an initial strengthening of comparative advantage in unskilled-labour-intensive manufactured products, such as textiles and clothing" (p. 27).

The objective of chapter 3 is to provide an overview of the trends and structural changes evident in China's economic growth since 1949. Statistics are presented in support of the argument that China's comparative advantage in agricultural products has

been declining for a long time. On the other hand, it is shown that comparative advantage in textiles and clothing has been strengthening. The comparative advantage in "other manufactures" shows no trend. These emerging trends are important not only for China but also for the rest of the world because, as Anderson's data shows, China is becoming more and more trade oriented.

The objective of chapter 4 is to examine the question of whether or not the rapid economic growth China experienced in the 1980s will continue. The author argues that "notwithstanding the economic and political instability of 1988-89, there are good reasons to assume not only that it will (continue to grow), but also that the relative openness of China's economy will continue" (p. 55). If current policies remain intact, then Anderson sees the role of agriculture declining in China's economy. "By the year 2000, farms will account for probably less than one quarter of GDP and half of all jobs, compared with around 30 percent and 60 percent, respectively, in 1987" (p. 76).

Chapter 5 speculates on how policy makers in China might respond to the loss of its comparative advantage in agriculture. He discusses price policy, foreign exchange policy, and policies affecting interregional trade. The overall conclusion of the chapter is that "the nonagricultural policy changes discussed here do little to raise one's hopes for China improving its agricultural comparative advantage in the longer run. If anything they tend to reinforce the likelihood that China will become increasingly more dependent on agricultural imports in the absence of intervention by the Chinese government to restrict such growth" (p. 94).

The purpose of chapter 6 is to review some of Anderson's previous (in his work with Rod Tyers) projections of China's future level of food self-sufficiency. In addition, this chapter evaluates the effects of China's economic growth on domestic and international markets for fibers. The overall conclusion is that "it is likely China's self sufficiency in agricultural products, particularly animal feedstuffs and natural fibres, will decline during the 1990s and beyond" (p. 107). Anderson suggests the impact of heightened Chinese imports on the international markets may be small because of the excess world capacity for food production, particularly in the United States.

To summarize, the overall theme of Anderson's book is that China's agricultural comparative advantage is on a downward trend. This is similar to the trend in Japanese agriculture described by Hayami in a book published a few years ago. While the falling comparative advantage theory may be an interesting hypothesis, Anderson's book falls short of rigorously

demonstrating this decline in the case of China. Nonetheless, the book is an ambitious and credible piece of work.

Colin A. Carter
University of California, Davis

Reference

Hayami, Yujiro. *Japanese Agriculture Under Siege: The Political Economy of Agricultural Policies*. New York: St. Martin's Press, 1988.

Bouman, F. J. A. *Small, Short and Unsecured: Informal Rural Finance in India*. Delhi: Oxford University Press, 1989, xi + 145 pp., Rs 120.

Bouman presents a lucid description of rural finance in India. He chose the district of Sangli in the state of Maharashtra as the battleground for his research. The choice of Sangli is aimed at providing fresh insights into the dynamic relationship between finance and development in an economy that is "in a rapid transition from a subsistence to commercial orientation" (p. 2). This is one area where the book differs from the existing literature dealing with the same subject matter.

The main topics of the book are presented in a logical manner. The book starts with a general review of the debate between formal and informal finance. This is followed by a discussion of the Indian rural finance policy. It then proceeds with a description of the sugar boom in Maharashtra. The discussion of the rural financial markets in Sangli starts with a general description of the various financial agents, and then focuses on key players in the informal financial markets, such as the Urban Credit Societies (UCS), the *bishi* (ROSCA/RESCA), pawnbrokers, and milk collectors. The last chapter gives an evaluation of the relation between finance and development.

Evidence of the systematic bias of policy makers against informal finance in India are sprinkled in several chapters. Examples of these are the Moneylenders' Acts, the Debt Relief Acts, and the lack of appreciation by policy makers of voluminous studies on informal finance in India done by noted researchers. The discussion about the Indian rural policy looks familiar to those who come from low income countries. The policy consists basically of creating highly specialized financial institutions that give less emphasis to savings mobilization and concentrate instead on the delivery of cheap credit to specific sectors of the economy. An example is the Primary Agricultural Credit Societies (PACS) which "operate as retail agents of government loans for agriculturists" (p. 39). The result was the rapid proliferation of unviable financial institutions with some equity implications because "wealthier farms are often the main defaulters" (p. 115).

Against this background is the dynamism of the

informal financial agents which consist of less regulated (e.g., the UCS and licensed pawnbrokers) and completely unregulated agents (e.g., the *bishi*, unlicensed pawnbrokers, and milk collectors). Bouman reports they have proliferated even in areas where numerous formal financial institutions already exist. Their emergence can be attributed mainly to the failure of formal financial systems to respond to the growing and changing demands of people for financial services. For instance, beneficiaries of the sugar boom who "had money to spend and were looking for new production and consumption alternatives" (p. 4) formed their own UCS because "the existing financial institutions could not accommodate them" (p. 4). *Bishis* became popular because of "the high return they offer on savings and the relatively easy access they provide to credit" (p. 56). The milk collectors who became informal lenders supplied loans to less affluent farms who could not obtain loans from the formal financial institutions, thus enabling them to respond to investment opportunities brought about by the dairy boom.

The informal sector in Sangli proved to be resilient and innovative. The collapse of several UCS caused by political intrusions and fraud committed by some of their officers elicited some positive responses, such as appointing motivated professionals with experience in managing a financial institution. Old forms of financial arrangements became obsolete and were replaced by new ones in response to the changing requirements of the rapidly growing economy. A case in point is the gradual disappearance of rotating savings and credit associations (ROSCAs) and the emergence of RESCAs (regular savings and credit associations) in Sangli. One weakness of the former is that only one player can be accommodated at one time; and, even then, he "cannot be sure of the moment he will receive the fund" (p. 55) because of the lottery mechanism. In contrast, several borrowers can simultaneously be accommodated in RESCAs. Interestingly, the recently created *bishis* tried to deal with the problem of shortage of capital by collecting "large-sized one-time lump sum payments" (p. 67) as shares of members and also lengthened their life cycle so that they can provide larger and longer loans.

The last section in the book deals with some nagging questions that have been raised since the time when informal finance caught the attention of researchers and policy makers. One such issue is whether formal and informal finance are substitutes or complements. Bouman has pointed out that in a penny economy, small and short loans are widely demanded. "It is precisely this small-loan specialization which precludes the substitution of the informal lender with a supposedly more public-spirited formal institution" (p. 125). Another issue is the linkage between formal and informal financial institutions. He cautions those who are trying to forge such linkage through some policy design. According to him, "the Sangli experience demonstrates that linkages seem to develop spontaneously whenever and wherever peo-

ple perceive such links as advantageous. Without such advantages, any policy of link-promotion will be an uphill struggle" (p. 127).

As one goes through the book, he will be bothered by a number of issues. For instance, the author did not discuss how the sample financial agents included in the study were selected. This is important to dispel any doubt about the positive contributions of informal finance in the economy. While the author mentioned some cases in which some of his Indian friends were trying to prove the absence of relationship between informal finance and economic development by citing "hearsay scandals of bishis, insisting that bishi-loans stimulated alcohol abuse" (p. 120), he could be accused of exaggerating the existence of such relationship by conveniently selecting the sample; hence, the need to discuss the sampling procedure. Mention should also be made of the instruments used in gathering primary data, i.e., whether structured interview schedules were administered to the respondents.

The author has pointed out that the supply of formal and informal loans in Sangli has been growing substantially with economic development and indicated that there is sufficient competition in the market by just looking at the number of financial agents. Yet, he did not provide an analysis of the behavior of lending rates over time. One would reasonably expect that real interest rates have gone down over the years if competition really did improve. In fact, the author did not provide a good analysis of the structure of lending rates. This is most glaring in this analysis of *bishis*.

Despite these questions, the book provides valuable insights on rural finance. Policy makers in LDCs should find this book useful in their work. For students of finance and development in LDCs who have been exposed to the elegant models of financial institutions' behavior, the book is a good eye-opener.

Mario B. Lamberte
Philippines Institute of Development Studies

Drèze, Jean, and Amartya Sen. *Hunger and Public Action*. Oxford: Clarendon Press, 1989, xviii + 373 pp., \$45.00, \$15.95 paper.

For those concerned with the alleviation of hunger in the developing nations through public action, this is a most important book. Hunger is defined as including famines and chronic undernutrition. Public action includes activities of both the state and individual members of society.

This is a big and challenging topic, but the authors are more than up to it. Drèze, a former lecturer in development economics at the London School of Economics, has extensive field experience in India. Sen, formerly Drummond Professor of Political Economy at Oxford and now Lamont University Professor at Harvard, is a brilliant polymath who has had a long-standing concern with hunger. They make a

formidable team. Their work was sponsored by the World Institute for Development Economics Research (WIDER).

The book builds on previous writings by both authors, particularly Sen's well-known *Poverty and Famines*. A companion three-volume set of readings edited by the authors, titled *The Political Economy of Hunger*, is being issued in 1991 by the same publisher. They have also done their share of reading: their bibliography extends some 77 pages and includes roughly 1,700 citations!

The text of the book is organized into four main parts. Part 1 provides background on "Hunger in the Modern World." Part 2, "Famines," reviews strategic and technical issues of famine prevention in India and sub-Saharan Africa. Part 3 is titled "Undernutrition and Deprivation" but is much broader in scope and might equally well have been called "Public Support Measures and the Quality of Life." Part 4 is a one-chapter summary emphasizing public action.

Overall, the authors state that "the primary focus of the book is on action, rather than on measurement" (p. vii). Hence, they have tried to make the discussion nontechnical and accessible. They have emphasized what can be done inside the country rather than through external aid. Public action of both a collaborative and adversarial nature is advocated.

They take a broad social security approach which separates the problems into two interrelated camps; those that can be handled through protection (famine), and those that can be handled through promotion (undernutrition). They view famine as a simpler and less pervasive problem than undernutrition; far fewer lives are lost to famine than to undernutrition. They see wages for the employable through public works programs, supplemented by relief for the unemployable, as the most promising way to tackle famine. In the case of undernutrition, they examine both growth-led and support-led security and provide a number of country examples of the latter.

The strengths of the book are many. It is well informed and well reasoned. It admirably combines great erudition with case studies and examples. It suggests, without "excessive shyness" (p. vii), practical courses of action. Everywhere it turns, it sheds new light and new insights. It is well organized, well written, and superbly produced (I did not spot any typographical errors). In short, it delivers, and in fine style.

Weaknesses are relatively few and may differ with the eye of the beholder. I think that the role of war and revolution in setting the stage for famine might have been given greater attention (brief mention on pp. 274-75). Also, the geographic and historical scope of the coverage of famines might have been broadened. While the alleviation of undernutrition is a much broader subject, I cannot help but feel that the authors went a bit overboard in most of part 3; undernutrition, as such, is given relatively little attention. Furthermore, the authors completely neglect the critical role that public agricultural research plays in expanding food supplies and lowering the per unit costs

of food production and consumption. Little is said, for either famine or undernutrition, about what can be done when governments are weak or chaotic, or of what the budgetary costs are likely to be (the latter subject is only briefly mentioned on pp. 251–52, 270).

But if some of these aspects might be more fully developed, it must be acknowledged that the authors and the subject have come a long way since Sen's 1981 book. The discussion that it engendered (especially in *Food Policy* since 1986) and further empirical work are reflected in a much better balanced and more informed discussion. The key role of production, for example, is given increased attention. This volume, in turn, may launch a new cycle of discussion and learning.

The result is a book that everyone concerned with hunger can read with profit. It will also help broaden and improve the vision of more general development specialists. It should, in short, make a major intellectual contribution to more enlightened public action to alleviate hunger in developing nations. It could, if read and heeded, change untold lives.

Dana G. Dalrymple

U.S. Department of Agriculture and Agency for International Development

Reference

Sen, Amartya. *Poverty and Famines*. Oxford: Clarendon Press, 1981.

Eicher, Carl M., and John M. Staatz, eds. *Agricultural Development in the Third World*, 2nd ed. Baltimore MD: John Hopkins University Press, 1990, x + 550 pp., \$55.00, \$18.95 paper.

The first edition of this book was published in 1984 and received generally good reviews by Carruthers and Roumasset, among others. The publisher states on the back cover of the second edition that, "Twenty of the book's thirty-two chapters are new, treating such topics as the role of women in rural development and policy reform in Latin America, Africa, China, and the Soviet Union. The readings address two basic questions about Third World agriculture: what is the strategic place of agriculture in national development, and how can the development of the rural sector be accelerated?" Net changes to the book are significant and worthwhile, though several "new" chapters are by different authors on the same topics.

"This book is designed to be used as a text in beginning graduate courses in international agricultural development and as a supplementary text in general economic development courses" (p. x). I have used the first edition in a course on the economics of agricultural development and have found it to be a generally excellent selection of the existing literature, although not all desired topics were covered.

Interestingly, outside materials on food price policy, global food balances, and productivity by farm size, which I have used to complement the first edition, are included in the second edition in their entirety or as portions of chapters. In my view, these additions make the book even more useful as a text for a beginning graduate course on agricultural development.

The second edition is divided into five parts that are organized in much the same way as for the first edition. Part 1, an overview, has an excellent review by the editors of the evolution of theories and development strategies for agricultural development since 1950. New sections include a discussion of the development themes and empirical evidence introduced in the 1980s and lessons and insights for the 1990s. This chapter, like the introductions to each of the four remaining parts of the book, provides a good background, overview, and some synthesis of the concepts covered in the following chapters.

Part 2, "Historical and Theoretical Perspectives," now includes four chapters, as compared with the seven chapters in the previous edition. Material on community development and the political economy of Latin America as well as two comment chapters have been deleted. New material includes two chapters on agriculture in transition and its shift toward industrialization, both positive additions. The induced institutional model chapter, by Ruttan and Hayami, remains in the new edition (pp. 97–112). However, none of the new institutional economics literature has been incorporated into the materials.

Part 3, "Food and Agriculture Policy" (formerly "Policy Analysis in a General Equilibrium Framework") includes only four of the nine chapters found in the first edition. The new title for this part more adequately reflects the material included in both editions, and rightly so, since previous reviewers pointed out the lack of any equilibrium framework in the first edition. The five new chapters cover some of the same subjects but with much more relevance and usefulness for development economists and practitioners. Particularly important are the chapters by Schuh, discussing the new macroeconomic situation (pp. 140–53), and by Bates, covering the political framework for decision making (pp. 154–59). These add important dimensions and insights to improving our understanding of the context within which agricultural development and policy change must take place.

Part 4, "Transforming Agriculture and the Rural Economy," has not changed significantly in terms of subjects covered but, again, the net change is very positive. The important topic of women's role in development has been added, not as a formality, but in recognition of its importance for effective policy development and implementation. The chapter, "Reflections on Land Reform and Farm Size," by Binswanger and Elgin (pp. 342–54), also is a significant replacement for two previous chapters in this section. This also can be said for the chapter on rural financial markets by Adams and Vogel (pp. 355–70).

Finally, part 5, "Case Studies of Economic Policy

Reforms," highlights the recent experiences of Latin America (new chapter), China (updated by new authors), and the Soviet Union (new chapter) in implementing policy reforms. These should be of special interest to readers watching the very recent attempts to change economic policies in Eastern Europe. In the last two chapters, two well-informed Africanists (Eicher and Lele) provide useful insights on the challenges facing most African nations overcoming their agrarian stagnation (pp. 503–39).

Eicher and Staatz have done a commendable job of drawing relevant materials from the extensive agricultural and economic development literature. The second edition includes many chapters focusing on current policy and development issues. It would be very difficult to find a better collection of materials in one book to use as a primary text for a beginning graduate course in agricultural development or as a supplemental text for economic development courses. The main criticism of the revised book likely to emerge is that it reflects the current development community emphasis on private enterprise, privatization, policy reforms, new technology, and "setting prices right." Most of the radical or opposing views of the neo-Marxists, dependency theorists, and others are largely muted in the second edition.

Ronald L. Tinnermeier
Colorado State University

References

- Carruthers, Ian. "Book Reviews." *J. Agr. Econ.* 37(1986):110.
Roumasset, James. "Book Reviews." *J. Develop. Econ.* 19(1985):201–10.

Fleischer, Beverly, *Agricultural Risk Management*, Boulder CO: Lynne Rienner Publishers, 1990, xv + 149 pp., \$25.00.

Smidts, A. *Decision Making Under Risk*. Wageningen Economic Series No. 18. Wageningen, Holland: Agricultural University, 1990, 329 pp., price unknown.

The literature on risk analysis and management in agriculture since the late 1960s has been accumulating at an increasing rate. Benchmarks of this growth are contained in the overviews on the subject by Dillon; Anderson, Dillon, and Hardaker; and Barry. We need another benchmark of the risk literature that has been accumulating since Barry's book, for the most part in journals and conference proceedings. The book by Fleischer and the monograph by Smidts contribute each in its own way to satisfying this need, Fleischer for the nonspecialist and Smidts for the specialist.

The purpose of Fleischer's book is "to provide nonspecialists with a guide . . . to risk and risk management . . . instrumental in understanding . . . agricultural policy" (p. xiii). It can be characterized

by its nontechnical approach to the discussion of risk management and its primary focus on the relationship between risk management and agricultural policy. Fleischer first surveys the subject of risk and how managers respond to risk, enumerates risk management tools, indicating why each is used and its effects on risk management strategies, and finally relates these to government programs and policies. The fact that all this is covered in just 149 pages indicates the brevity with which topics are covered. This book is clearly intended for those with no prior knowledge of the subject of risk and its management.

The monograph by Smidt reports his thesis on the application of a "relative risk attitude" model to potato-marketing behavior of producers.¹ It is a very thorough technical treatment of several aspects of risk in decision making associated with a real-world situation. After describing the decision problem confronting farmers, Smidt reviews and evaluates the various theories used in decision making under risk, carefully integrates these theories into the relative risk attitude model, thoroughly assesses the methodological problems of measuring risk perceptions and assessing risk attitudes, and finally analyzes the results of applying his model and measurement tools to a multiyear marketing strategy. This work will be valuable to anyone contemplating the analysis of risk or risk management in agriculture and to those wanting more than a cursory introduction to the area.

While Fleischer's book is useful for the purpose for which it was written, probably too much was attempted in the limited space allotted for each topic. In the latter chapters, the treatment of some common tools for managing risk in agriculture, forward pricing, and the role of government commodity programs are well done and should be informative to the intended audience. The first two parts are structurally less sound. A book of this intent should bring some order out of the profusion of information on the subject that is available and present it in clear, concise, and unequivocal terms. This book does not do this. For this reason, and the notable lack of citations, the book will have limited use as a course text.

The review and evaluation of theoretical and risk attitude measurement concepts by Smidt are very well done. The treatment of both topics goes beyond that normally expected in a thesis. With a strong focus on a specific applications problem, Smidts critiques and evaluates relevant theory with considerable thoroughness, providing a sound basis for the hybrid approach used in his model. He also is as thorough in his development and testing of alternative numerical methods for applying his theoretical constructs to a survey of 250 producers over a two-year period. The final chapter on evaluation and conclusions by itself is well worth reading and along with the chapters on

¹ "Relative risk attitude" is a theoretical bridge between Bernoulli's "strength of preference" approach and von Neuman-Morgenstern's "indifference utility" approach first introduced by Bell and Raiffa in 1982.

theory and methodological development provide a great deal of useful information. Those involved in studying more than just the theory of risk management and analysis will find this work to be helpful.

Although neither of these works completely satisfy the need for a benchmark book on risk, both in their own way do some of the job until one comes along. In particular, both have excellent bibliographies that capture the literature as it currently exists.

Walter L. Fishel
Ohio State University

References

- Anderson, J. R., J. L. Dillon, and B. Hardaker. *Agricultural Decision Analysis*. Ames: Iowa State University Press, 1977.
- Barry, P. J., ed. *Risk Management in Agriculture*. Ames: Iowa State University Press, 1984.
- Bell, D. E., and H. Raiffa. "Marginal Value and Intrinsic Risk Aversion." *Decision Making, Descriptive, Normative and Prescriptive Interactions*, ed. D. E. Bell, H. Raiffa, and A. Tversky, pp. 384-97. Cambridge: Cambridge University Press, 1988.
- Dillon, J. L. "An Expository Review of Bernoullian Decision Theory in Agriculture: Is Utility Futility?" *Rev. Mktg. and Agr. Econ.* 39(1971):3-80.
- Gray, Kenneth R., ed. *Soviet Agriculture: Comparative Perspectives*. Ames: Iowa State University Press, 1990, xiii + 284 pp., \$29.95.
- Wädekin, Karl-Eugen, ed. *Communist Agriculture: Farming in the Soviet Union and Eastern Europe*. London and New York: Routledge, 1990, xviii + 331 pp., \$112.00.
- Both of these books deal with Soviet agriculture and agriculture in other selected Eastern European countries with centrally planned economics. Each book is comprised of papers from two separate conferences dealing with Soviet and East European agriculture and agrarian affairs. The contributors to these books represent several disciplines including economics, history, geography, and political science and exhibit considerable competence in the technical, social, economic, and political aspects of agriculture and agrarian affairs in the Soviet Union and selected Eastern European countries. The purpose of both books is to represent the types of analyses needed to understand the changes in agriculture and agrarian affairs in the Soviet Union and some Eastern European countries which led to the type of agriculture these countries had in the late 1980s.
- The book edited by Gray is arranged in two parts. The first part is concerned with the organization and performance of Soviet agriculture. There are nine chapters in this first part, dealing with topics such as agricultural productivity trends in the Soviet Union compared with other centrally planned economies, costs of agricultural growth and development in the USSR, Soviet food imbalances, Soviet utilization of meat and dairy processing capacities, agricultural pricing and policy under Gorbachev, recent developments in the agricultural collective contract, recent changes in Soviet rural housing policy, and trends in Soviet dryland farming and soil conservation practices.
- The second part of this book consists of five chapters dealing with technologies and information which could be drawn upon by Soviet agriculture. In the first chapter of this part, chapter 10, the Soviet scientific and technical information system concerning agriculture is described and evaluated in terms of effectiveness. In the next chapter, the contents of Soviet writings on foreign agricultural experiences are analyzed for possible impacts on Soviet agricultural policies. The last three chapters deal with potential lessons for Soviet agriculture from the Bulgarian experience with centrally planned agriculture, private agriculture in other socialistic countries, and Hungarian and Chinese experiences. These latter three chapters provide some informative comparisons between Soviet agriculture and agriculture in other centrally planned economies.
- The book edited by Wädekin contains three parts. The first part, comprised of six chapters, considers the historical, environmental, and social background of agriculture in the Soviet Union. The eight chapters of the second part focus on the current agricultural and industrial policies in the Soviet Union and their impacts on agriculture. The last seven chapters of the book comprise the third part and deal with agricultural experiences in other countries with centrally planned economies and the implications of these experiences for Soviet agriculture.
- The six chapters which comprise the first part of the book edited by Wädekin address issues such as the role of agriculture in Soviet industrialization, land-use zones and soil degradation, technical efficiency in cotton production, new attitudes in rural areas, impact of the Chernobyl incident on the Soviet food supply, and structural contrasts and convergence in socialist and capitalist agriculture. These six chapters provide an excellent background for understanding Soviet agriculture. The next eight chapters analyze recent and current agricultural policies such as *perestroika* in agriculture, the March 1986 agricultural decree, uncertainties over reform, normative planning in Soviet agriculture, Raion agroindustrial associations in Soviet agriculture, organizational innovation in agricultural work, economies of small scale in Soviet agriculture, and private plots as system stabilizers. The third part of the book consists of seven chapters which are concerned with reforms in socialistic agriculture in Bulgaria, Poland, German Democratic Republic, and Hungary. In each of these latter chapters, attention is devoted to an analysis of major reforms in agriculture in the four selected Eastern European countries.
- These two books are useful references not only to

those interested in understanding agriculture and rural affairs in the Soviet Union and other Eastern European countries with centrally planned economies but also to those concerned with reforming or dismantling socialist agriculture in other parts of the world. These two books provide comprehensive analyses and information about various agricultural programs and policies in the Soviet Union and other centrally planned economies. Specific programs of various agricultural plans are analyzed and evaluated in terms of their physical, economic, social, and political consequences and their degrees of success or failure. The authors of the chapters in these books address why various agricultural programs did not work any better than they did and why many did not function as planned. These two books should be of interest to historians, political scientists, economists, and other social scientists concerned with socialistic agriculture and its role in planned economies.

Joseph Havlicek, Jr.
Ohio State University

Jackson, John H. *Restructuring the GATT System*. New York: Council on Foreign Relations Press, 1990, vi + 121 pp., \$14.95.

The importance of the General Agreement on Tariffs and Trade (GATT) for world agriculture has become more evident as a result of the recently completed Uruguay Round. As agriculture becomes more integrated into the GATT over time, agricultural trade economists will increasingly need to understand the history and inner workings of the GATT. This book, written by one of the leading legal scholars on international trade, serves as an excellent introduction to the topic.

As suggested by the title, the author espouses change in the GATT "constitutional" system. In the final chapter, he outlines a proposed charter for a World Trade Organization, a broad international commerce organization which would "serve and 'shelter'" other treaty instruments, such as the GATT. Its purpose would be similar to the non-GATT portions of the International Trade Organization (ITO) and to the Organization for Trade Cooperation (OTC), international organizations negotiated in the early post-World War II period but rejected by the U.S. Congress. Jackson's proposal should be taken seriously and warrants further attention. However, for most readers, the importance of the book will be its thorough and concise description of the GATT system as it has developed and now stands.

After an introduction, Jackson begins part 1, "The Defective Constitution of GATT" with chapter 2 on "The History and Perspective of GATT." This chapter examines the wartime and early postwar environment which led to the negotiation of a triumvirate of international economic organizations: the International Monetary Fund, the World Bank, and the ITO. Drafting of the ITO was completed in 1948 in Ha-

vana. The Havana Charter included agreements on economic development (including the regulation of foreign investment), restrictive business practices by private as well as public business (an internationally applied antitrust policy), commodity agreements and, notably, a well-defined organizational structure. However, the ITO was never approved by the U.S. Congress; and the GATT, designed to be subordinate to the ITO, continues to be applied provisionally. While Jackson describes the GATT as having been "much more successful than one could fairly predict in 1951 when the ITO idea died," he argues that the current constitutional structure of the GATT and its provisional application (which, for example, allows national policies predating the GATT to continue in force because of a grandfather clause) inhibited GATT's ability to deal with troubled industries such as textiles, agriculture, and steel.

The GATT "constitutional" structure itself is the topic of chapter 3. Here Jackson addresses the questions of GATT membership, voting rules, amendments, side codes (such as the International Dairy Arrangement), and the relationship of GATT and national legal systems. While these are potentially dry topics for economists, Jackson sprinkles his writing with examples and historical precedents which indicate the importance of these issues and expose the frailty of the current system. A short chapter on the history of GATT negotiation rounds, with special emphasis on the Uruguay Round, completes part 1 of the book.

In part 2, "Perspectives on International Economic Institutions," Jackson begins by discussing "The Jurisprudence and Policy of a GATT Constitution" (chap. 5). Some of the institutional problems of GATT that Jackson perceives include (a) the treaty structure, which remains provisional and allows some parties to question whether they are bound by GATT rules; (b) the amendment provisions, which make amendments exceedingly difficult and limit their application to countries which accept them; (c) a lack of clarity in the relationship of "side codes" to the GATT; and (c) the dispute settlement process.

The last of these is the major topic of chapter 6, "Rule Application and Dispute Settlement." Jackson notes disagreement over whether the goal of GATT dispute settlement procedures should be to solve the instant dispute (by conciliation, obfuscation, power-threats, or otherwise), or to promote certain longer-term goals. In making his case for adopting the latter objectives, Jackson carefully reviews the evolution and current state of dispute settlement under GATT. Complaints about the procedures are numerous: (a) they are slow; (b) a loser may block adoption of the resulting panel report; (c) it is sometimes unclear whether the GATT procedure or another procedure outlined in a "side code" is the appropriate one to follow; and (d) governments are often slow to implement any required changes. It is interesting that most cases are brought by industrial countries against other industrial countries; however, the markets in

question are often not industrial. Jackson notes that, especially with respect to the European Community (EC), agricultural products are often involved in disputes. The EC is virtually alone in its opposition to attempts to strengthen the GATT dispute settlement procedure, such as one outlined by the author. Jackson gives compelling reasons why neither disputant should have the right to block adoption of a panel report, as they currently can.

A particular future challenge for the GATT involves the integration of large centrally planned economies such as China and the Soviet Union. In chapter 7 the author argues that these countries should be allowed into the GATT (or some other world trading organization), noting the prevention of war was an objective of the postwar international economic system. This "universal approach" to membership will cause particular problems and require the development of an "interface mechanism." Arguing that it is not the purpose or role of the GATT to apply pressures on sovereign nations to accept market-oriented economic principles, he writes that interface rules should be designed under which the applicant non-market economy can trade with the existing GATT market economies to the most feasible extent.

The book concludes with a detailed plan for "Reforming the GATT System" in chapter 8. The charter for Jackson's World Trade Organization (WTO) would build on the existing GATT by providing an institutional structure for the GATT and other agreements while serving the institutional needs for service sector agreements. It would also implement new agreements (regarding, for example, intellectual property rights) resulting from negotiations under GATT, establish a centralized panel procedure for dispute settlement under GATT and the "side codes," and make explicit the relationship of other treaty instruments to the WTO.

Jackson's book will be of interest to economists and agricultural economists who want to learn more about the GATT and the international trading system. Although his proposal for reform in the last chapter is sketchy, he pinpoints many problems with the current system, and his proposal should provoke much thought and research.

Bradley J. McDonald
Economic Research Service,
U.S. Department of Agriculture

Lobao, Linda M. *Locality and Inequality: Farm and Industry Structure and Socioeconomic Conditions*. Albany: State University of New York Press, 1990, xiv + 291 pp., \$16.95

Lobao takes on a massive task as her "book explores how the organization of economic production in farming and industry generates socioeconomic inequality across different localities in the United States" (p. 1). As such, it represents a significant contribution to an aging literature on farm structure and com-

munity relationships. The book would contribute to a graduate course on farm structure or rural development. It is also important reading for scholars in these areas.

The major contribution rests with a challenge to traditional work following Goldschmidt, who tested the hypothesis that communities influenced by small farms are better bastions of democracy with improved quality-of-life than are communities influenced by large farms. Goldschmidt used a case study of two communities in California to support this hypothesis. Lobao builds on work for the Office of Technology Assessment (see Swanson) to examine the Goldschmidt hypothesis.

Lobao devotes three chapters (2-4) to reviewing literature and building toward a theoretical framework for her empirical tests. Her critique and integration of many conflicting paradigms is particularly noteworthy. I found these chapters informative and well-developed. She reviews both the sociology and agricultural economics literature, bringing the reader up-to-date on the studies performed in this important area.

A major and significant theme in the theoretical development is that rural communities are influenced by many forces other than farming. Lobao elaborates: "I develop a conceptual framework that ties farming and industry structure and other structural factors to inequalities in socioeconomic conditions across localities. This involves the interplay of three sets of forces: economic structure, spatial or locational structure, and human agency in the form of characteristics that empower workers and the households. My perspective is developed through a synthesis of political economy and other structural approaches" (p. 5).

In chapter 5, she presents data that is subsequently used for analysis. Of note to the agricultural economist is the development of an index of farm structure. Building on work of Wimberley, Lobao develops indexes that capture differences "in farm organization, scale, and operator characteristics at the county level" (p. 98). Three patterns of farming are developed: (a) industrialized farming, (b) larger family farming, and (c) smaller family farming. Farm structure indexes are used with a variety of county data on industry structure, labor, spatial characteristics, and measures for inequality (e.g., median family income, poverty, income inequality, unemployment, infant mortality, etc.).

Chapter 6 presents national models, and chapter 7 presents regional models used to test hypotheses regarding community inequality. Because of data availability, the analysis is restricted to census base periods around 1970 and 1980. In addition to cross-sectional models, lagged models are developed to test how 1970 conditions influence 1980 measures of inequality. I was left wondering why the author did not expand these models to reflect some recent work using similar data (see Swanson). In this work the change between the two periods is modeled by taking the dif-

ference in both independent and dependent variables for the two points in time.

Generally, Lobao has numerous findings with regard to "locality and inequality." I shall focus only on the farm structure findings. An underlying theme of the book is that assumptions of a purely linear relationship between farm structure and community well-being or inequality should be challenged. Lobao recognizes work by Skees and Swanson that influenced her thinking in this regard. She summarizes her farm structure findings as follows: "This study provides limited support for the Goldschmidt hypothesis that industrialized farming reduces socioeconomic conditions across U.S. counties. But the detrimental impacts of this production system are neither automatic nor strong; further, they are not apparent at any particular historical point (cross-sectionally) in every region of the country This study also has shown that the effects of industrialized local farming are not predetermined but can be modified by nonfarm factors" (p. 213). She goes on to state "the findings with regard to smaller family farming indicate that the Goldschmidt hypothesis needs serious revision" (p. 214). In short, her findings demonstrate negative relationships between well-being measures and small and industrialized farming. "A consistent finding of this study is that a production system of family operated commercially-oriented farms—larger family farming—results in better socioeconomic conditions for localities" (p. 216).

At times, Lobao slips into a common trap of making too much of her model relationships as she leads the reader into thinking that cause and effect are being presented. However, in her conclusions she is very clear about this shortcoming of many of her models. She is to be commended for a massive undertaking and a significant contribution to a literature that has been lacking in its development.

Jerry R. Skees
University of Kentucky

References

- Goldschmidt, Walter. "Small Business and the Community: A Study in the Central Valley of California on Effects of Scale of Farm Operations." Hearings before the Subcommittee on Monopoly of the Select Committee on Small Business, U.S. Senate, 90th Congress, 2nd sess., Washington DC, 1968.
- Skees, Jerry R., and Louis E. Swanson. "Farm Structure and Rural Well-Being in the South." *Agriculture and Community Change in the U.S.: The Congressional Research Reports*, ed. Louis E. Swanson, pp. 238–321. Boulder CO: Westview Press, 1988.
- Swanson, Louis, E., ed. *Agriculture and Community Change in the U.S.: The Congressional Research Reports*. Boulder CO: Westview Press, 1988.
- Wimberley, Ronald C. "Dimensions of U.S. Agristucture: 1969–1982." *Rural Sociology* 52(1987):445–61.
- Longworth, John W., ed., *China's Rural Development Miracle, with International Comparisons*. Queensland: University of Queensland Press, 1989, 450 pp. plus list of participants, price unknown.

This book is a collector's item. For almost six decades, the International Association of Agricultural Economists (IAAE) endeavored to arrange a professional conference in China. The preface to this book reports that the Chinese scholar C. C. Chang first proposed a China Conference during his visits with IAAE members in Europe in 1933. Chang's original travel objective of attending the IAAE meeting that year was frustrated because of its abrupt cancellation as a result of problems in the host country, Germany. Chinese participation in IAAE meetings was limited during the ensuing years. The conference finally became a reality in October 1987 when agricultural economists from fifteen countries gathered near Beijing for professional discussions with Chinese colleagues. This book is a record of the symposium, the theme of which was "Rural Development Strategies: Theory and Practice." Appropriately enough, an active participant in the symposium was C. C. Chang.

The symposium organizers planned for the presentation of fifty papers, half to be contributed by IAAE members and half by members of the Chinese Association of Agricultural Economists. The aim was for foreign participants to explain other country experiences with rural development. Likewise, the Chinese were to inform visitors about China's rural development. Most of the papers were presented in small groups with adequate time for translation and discussion. The small group approach encouraged frank discussion and exchange despite language and cultural barriers.

The book consists of an eight-page editor's introduction, fifty-one chapters, of which thirty-four are full papers and seventeen are one- or two-page abstracts, and a list of participants. Approximately thirty papers concern China directly, while the remainder summarize economic and rural development experiences in other countries, sometimes with explicit implications of those experiences for China. The papers presented by Chinese authors tended to be descriptive, while the non-Chinese authors often times presented theoretical work with some quantitative analysis. Only a few papers reported formal quantitative models.

Part I sets the stage for the book through three short papers written by Chinese authorities. These papers provide a brief rationale for the political and economic policies followed by China in the post-World War II period with particular emphasis on the landholding and agricultural marketing systems and a description of the key reforms associated with the household based contract responsibility system established in the late 1970s. The authors note the challenges currently faced by a country with more than a billion people, with less than one hectare of arable land per peasant household, and with a serious short-

age of physical infrastructure, institutions, and administrative ability.

Part 2 contains papers reporting on aspects of development experiences. The papers report the experiences of specific countries such as India, Korea, Japan, Eastern and Western Europe, and Israel. Other papers explore specific issues such as central place theory applied to Germany and the impact of trade and macroeconomic policies on agricultural growth as observed in several Latin American countries. A common theme linking many of these papers is the role of government policy in shaping the speed and pattern of the agricultural changes noted in the countries studied.

The third part contains fourteen papers under the general heading, "Recent Chinese Experience and Prospects for the Future." Common themes in these papers include past trends in agricultural production; average cereal supplies per capita; trends in population, labor force, and surplus agricultural labor; the impact of changes in policies beginning in 1980; projections of future supplies of and demand for food and agricultural products (usually without describing the methods used); and the need for more off-farm employment to absorb labor. One paper reported the results of analysis that showed substantial economic gains could be made through greater regional specialization in production based on comparative advantage.

One of China's serious shortcomings is poor infrastructure, especially transportation, which impedes greater regional specialization and interregional trade. Another problem concerns the future supply and use of chemical fertilizers. Two chapters in part 4 analyze infrastructure problems, and a third chapter discusses fertilizer supply and demand issues in comparison with the experience of India. Both countries have sharply expanded fertilizer use; but the methods to achieve this objective have been quite different, with China relying more heavily on public institutions to circumvent the shortcomings of market underdevelopment.

Four chapters in the next section deal with the subjects of rural investment and finance. They consider how agricultural investments have evolved and may be affected by the agricultural reforms that have sharply altered the urban-rural terms of trade. The control of investment funds has undergone a dramatic shift from central government to township and village enterprises. Enterprises and households now control larger amounts of gross domestic savings than the central government. The reforms made in the rural financial system to capture and allocate these savings are described.

Part 6 consists of seven papers grouped under the theme, "Ownership, Management, and Reward Systems." The papers examine the development of rural cooperatives, state farms, freshwater fish ponds, and the position and role of women in rural development. One paper compared the performance and incentive structure of Chinese and Canadian farms using the

entropy metaphor from systems theory. Another paper reports on an analysis of a sample of farms in Jiangsu Province designed to test the efficiency and equity implications of the household production responsibility system.

The final section, with eight papers, focuses on technology, research, and environmental issues. They cover the topics of technology transfer, research priority setting, farming systems research, information systems for microlevel decision making, and environmental issues associated with agricultural development. The paper on research priorities contains a preliminary analysis of the impact of agricultural commodity research in China, and discusses how different research objectives will alter the choice of commodities selected for primary research emphasis.

The papers in this volume present an insightful, albeit incomplete, view of the 1978 to 1986 developments that became known during the symposium as "China's Rural Development Miracle." The household-based contract responsibility system fundamentally altered producer incentives and sparked a rapid expansion in production, especially in the non-grain sectors. This pattern of decentralization of investment decisions is consistent with some of the rural development needs expressed in the symposium. First, off-farm employment opportunities will increase with the potential of absorbing some of the underemployed and unemployed rural labor. Second, increased local investments in rural towns and small cities may lead to a more socially desirable urbanization pattern than found in big cities in Western countries. But there are potential pitfalls as well. Pursuing regional self-sufficiency in the past in China is generally believed to have been enforced at a high economic cost. The poor state of infrastructure makes it difficult for China to now pursue the objective of specializing in regional comparative advantage under more liberal market conditions. Local decision makers who control more financial resources may choose to intensify growth and development in their respective regions and not invest in roads, ports, and communication facilities needed for expanded interregional trade. The development of financial markets to tap rural savings and facilitate the interregional flow of funds, improvements in tax collection and investments in public goods, and improved incentives for private infrastructure investment may determine if the short-term dramatic "development miracle" can be transformed into long-term sustainable growth.

John Longworth, Glenn Johnson, Theodor Dams, and Jim Hildreth of IAAE are to be congratulated for their persistent work with the Chinese authorities over several years to finally realize this symposium, and to help C. C. Chang achieve his dream. The papers in this volume are a tribute to their work. More important, the intense discussions held at the meeting and the circulation of the Chinese proceedings now underway in China may contribute in important unknown ways to the policy debates occurring in that country. We can only hope that the abrupt rupture in

international exchanges following the Tiananmen Square tragedy is only a temporary interruption of the exchanges between Chinese and foreign agricultural economists that began with so much promise in this symposium.

Richard L. Meyer
Ohio State University

Pearce, David, Edward Barbier, and Anil Markandya. *Sustainable Development: Economics and Environment in the Third World*. Brookfield VT: Edward Elgar, 1990, vii + 217 pp., \$48.95.

As the development fad of the early 1990s, "sustainable development" may or may not deserve the attention of serious scholars. It is clear, however, that this particular idea has caught on among politicians and those who administer development assistance. The question remains, however, whether or not serious work by economists (or others) will have any impact on what the development community—and host governments—do with respect to natural resources in the tropics. The answer to that question, I suggest, depends upon whether or not the academic community is able to make meaningful programmatic sense out of the concept of sustainable development. The title and stated purpose of this book suggest that general goal.

We find here six chapters of reworked consulting reports and accounts of field work by the authors, preceded by three chapters that attempt to define the conceptual context for sustainable development. I first comment briefly on the six "empirical" chapters.

The empirical substance of the book consists of reports that appear to have been written for other audiences. These reports deal with (a) upland watersheds in Java; (b) forests on the outer islands of Indonesia; (c) the Sudan; (d) Botswana; (e) Nepal; and (f) Amazonia. The material in these chapters is descriptive as well as prescriptive. That is, certain natural resource situations are described and statements are made that, unless sustainable development practices are adopted, there will be hard times ahead. Policies in each country are discussed for their implications for sustainable development. There are useful data in these chapters and nice descriptions of specific resource problems.

Readers with experience in any of these countries are not likely to find much here that is not already well known to them. Yet, for the newcomer to natural resource issues in the developing world, this material will be most helpful in providing a perspective on certain problems.

But if the idea of sustainable development is to have any substantive impact on development assistance, academics must offer up clear ideas about the operational content of the concept of sustainable development. There is an attempt in this regard. The first three chapters provide the "theoretical" context for

the empirical (descriptive) chapters that follow, and it is in these three chapters that the authors attempt a coherent treatment of the idea of sustainable development. Chapter 1 discusses "Sustainable Development: Ecology and Economic Progress." Chapter 2 is titled "Discounting the Future," and chapter 3 is concerned with "Economic Appraisal and the Natural Environment."

It is in chapter 1 that one hopes for real substantive progress on the concept of sustainable development. Here the authors note that, while the notion of sustainable development is prevalent in development thinking, "little headway appears to have been made in terms of a rigorous definition of the concept. Therefore, not surprisingly, efforts to 'operationalize' sustainable development and to show how it can be integrated into practical decision-making have been few and generally unpersuasive" (p. 1).

The authors start by defining development as a vector of desirable objectives or attributes which a society might seek to achieve or to maximize. Listed here are (a) increases in real income per capita, (b) improvements in health and nutritional status, (c) educational achievement, (d) access to resources, (e) a "fairer" distribution of income, and (f) increases in basic freedoms. Sustainable development is then, according to the authors, a situation in which the development vector does not decrease over time.

After defining sustainable development, an attempt is made to define the necessary conditions for achieving sustainable development. The authors note some confusion between these two aspects and hope to clarify matters by this important distinction. The key necessary condition for sustainable development is then said to be "constancy of the natural capital stock. . . . More strictly, the requirement is for non-negative change in the stock of natural resources and environmental quality. In basic terms, the environment should not be degraded further but improvements would be welcome" (p. 4).

The authors properly note that natural capital can be converted into "man-made capital," and that one feasible goal is to seek a constant level of total capital in an economy. In that way, oil that is extracted can be converted into other forms of capital that will keep on giving long after the oil is exhausted. I find this discussion in chapter 1 very helpful and suggestive. My only complaint is that the authors did not apply their considerable talents to a more substantial and elaborate treatment of the subject. This question is, after all, the very core of what is meant by sustainable development.

The final "conceptual" chapter (3) is concerned with the economic appraisal of projects and programs of relevance for natural resources. Sustainable development is assured, we are told, if the benefit-cost analysis of a constellation of projects (a program) is modified by:

. . . setting a constraint on the depletion and degradation of the stock of natural capital. Essentially,

the economic efficiency objective is modified to mean that all projects yielding net benefits should be undertaken subject to the requirement that environmental damage (i.e. natural capital depletion) should be zero or negative. However, applied at the level of each *project* such a requirement would be stultifying . . . At the *programme* level, however, the interpretation is more interesting. It amounts to saying that netted out across a set of projects (programme), the *sum* of individual damages should be zero or negative. (page 59)

This idea is operationalized by including in any set of projects a shadow project whose purpose is exactly to offset the detrimental effects of the proposed investment portfolio, not at the project level, but so as to hold the natural capital stock constant across all activities in the nation. While an interesting idea, note that it is at odds with the discussion in earlier chapters where it was recognized that the aggregate capital stock, consisting of both natural and man-made assets, is of overarching economic importance. Because this chapter is only three and one-half pages long (albeit with a five-page appendix), one wishes that more time and thought had been devoted to these critical issues.

For it is here that one encounters the difficulty in pondering sustainable development as it relates to natural resources and the environment. Some participants in the discussion will regard natural resources and the environment as meaningful objects in their own right. Others will regard the values which flow from natural resources and the environment as the objects of policy interest. Finally, there are those who will regard all capital—natural and man-made—as mere means to an end; that end being to give rise to income and wealth. The formulation adopted by the authors seems consistent with the second of these formulations. Their formulation ignores the intrinsic values of the natural environment, and it confines relevant economic benefits to those arising from natural resources.

In a sense, this book has two personalities. The first three chapters are suggestive but stand as incomplete if we are to develop an operational concept of sustainable development. I wish the authors had spent much more time here; they are well-qualified to make significant contributions in this regard. Had this more elaborate work been undertaken on the conceptual front, the six descriptive chapters could have been written very differently. One almost gets the impression that the latter chapters existed before the three early conceptual chapters and were then modified to give them a flavor of sustainable development. The concepts in the first three chapters are rarely present in the descriptive chapters, though sustainable development is of course mentioned often enough.

I have little quarrel with what is here. It is what is not here that is troublesome. Pearce, Barbier, and Markandya have much to offer us, and I hope they

do just that in a more substantial undertaking in the very near future.

Daniel W. Bromley
University of Wisconsin

van der Meer, Cornelis L. J., and Saburo Yamada. *Japanese Agriculture, A Comparative Economic Analysis*. London: Routledge, 1990, xvi + 217 pp., \$65.00.

This book is a welcome addition to several recent studies on Japanese agricultural problems. In the preface, the authors point out that Japan's "major problem is the very high prices of agricultural commodities in comparison with international prices." Hence, "through international comparative economic analysis, this study tries to reveal to what extent recent Japanese agriculture has stagnated and to analyze the reason why" (p. xiv).

The method used is total factor productivity analysis. After citing several studies on "productivity performance in modern Japanese agriculture," including the work of Yamada himself, the authors find that "the concepts of output and productivity used in these studies have shortcomings for inter-country comparison of real value added and economic efficiency" (p. 2). The purpose of this volume is then to fill the gap in the literature. "Our study is the first to make a detailed real comparison of total production, current inputs, use of fixed capital and value added in Japan and other countries" (p. 2). The authors summarize the finding of their work as follows. "It reveals that real net value added in Japan's agriculture at US and Dutch prices is negative. Our study also analyses the mechanisms by which this negative outcome has been attained: i.e. the dynamic interaction of over-protection of agriculture and excessive use of resources in this sector" (p. 2).

The book consists of eight chapters, four appendices, eight figures, and seventy-two tables. After a brief introduction, in chapter 2 the authors carefully compare the "growth rates of sectoral labor productivity (gross value added per worker) over the period 1960–85 for eighty-one countries grouped by major region and developing status" (p. 24). It appears that, among the seventeen industrial economies, Japan is the only country which has a growth rate of labor productivity smaller for agriculture than for industry (p. 26). This is the case especially during 1975–85. Using least squares method and correlation, the causes of differences in agricultural labor productivity growth are attributed to the growth of labor productivity in nonagricultural sectors, share of agricultural employment (p. 35), price level for final output (p. 38).

Country comparisons of the levels of labor and land productivities and their growth rates for years 1963, 1975, 1985 are presented in chapter 3. It "identified

stagnation of gross value added per worker in Japanese agriculture in the 1975–85 period, poor performance relative to the predicted level on the basis of international comparison, and a negative effect of the high price level on productivity performance based on the confirmed inverse U-shape hypothesis" (p. 67), which, incidentally, is a sort of backward-bending labor supply curve.

The book then makes an in-depth comparison of agricultural development in Japan and the Netherlands (chap. 4). "A remarkable finding" (p. 11) is that the gap in real agricultural productivity, whether measured in 1975 Japanese prices, or Dutch prices, or international dollars, has increased gradually since 1880 (p. 77), while the gap in land productivity decreased (p. 80). Furthermore, "compared to the corresponding Dutch total factor productivity performance, Japanese agriculture lagged much behind the Netherlands, in particular for the period 1975–85" (p. 89).

The comparison extends to the first half of chapter 5. "The productivity gap is biggest for arable production and smallest for horticulture" (p. 129). The former gap "supports the general belief that land scarcity and small farm-sizes are obstacles for achieving high productivity in Japan" (p. 129). However, these reasons are not enough to explain the gap in horticulture and livestock, which are not very land intensive (p. 129).

Two more countries, Taiwan and the United States, are then added for an in-depth comparison in the second half of chapter 5. The comparison of real labor productivities with Taiwan at different relative prices (p. 117) shows that "Taiwan has taken over the leadership role in productivity in Asian agriculture from Japan" (p. 118). It shows that "Japan's level of agricultural protection is among the highest in the world" (p. 131). This chapter is more on methodology of international comparison than country comparison. The authors compare their findings with other studies, and find that "growth rates at constant national

prices may differ considerably from growth rates at constant international prices" (p. 129).

The reasons of Japan's poor performance in agriculture are analyzed in chapter 7. The authors argue that rapid growth in industry lead to "strong protection for agriculture" (p. 135). The Japanese political structure is such that "the rural votes weigh much heavier than urban votes . . . that the continuously ruling Liberal Democratic Party depends much on rural votes and funding . . ." (p. 139). The concern of food security and weak consumer voices enable the government to maintain continued protection and high prices of agricultural goods.

The book concludes with brief discussions of the future (chap. 7) of Japanese agriculture, and the relevance of Japan's post-war development experience for developing countries (chap. 8).

The data used are mainly from international sources like FAO, APO, ILO, and CBS, and are carefully documented in extensive footnotes and appendices. Although technical terms are also explained in appendices, the book is not easy to read and comprehend. Nontechnical readers may lose the focus of some chapters as the materials tend to be repetitive, fragmentary, and sometimes extraneous to the title of the chapter. The quantitative tools used are elaborate classical economic statistics, correlation, and regression. We may note that, in recent years, there are many papers dealing with the U.S. productivity slowdown in manufacturing. The method used is also total factor productivity analysis. Thus, more econometrics-oriented readers may wonder whether the new approach can be applied to analyze Japan's productivity slowdown in agriculture. Until this new approach is done, either by the authors or by others, current researchers on Japan's agricultural development will undoubtedly find this book one of the most comprehensive and useful references.

Frank S. T. Hsiao
University of Colorado

mouton de gruyter

Berlin · New York

**europaean review
of agricultural economics**

Editor: Arie Oskam
Agricultural University
Wageningen, The Netherlands

The **European Review of Agricultural Economics** serves as a forum for discussions about the development of theoretical and applied agricultural economics research in Europe and for stimulating ideas regarding the economic problems of agriculture in Europe and other parts of the world.

The **ERAE** also promotes discussion on national resource use, protection of the environment, marketing of agricultural products and development of rural areas. Throughout, the **ERAE** strives for balanced coverage of all issues in agricultural economics: production economics, operations research and farm management problems, agricultural policy, including farm incomes and farm structure, regional planning and rural development, supply analysis, factor markets, demand analysis and marketing, international trade and development, statistical and econometric methods, etc. Original articles as well as full or abstracted articles which have already appeared in national publications and/or in other languages are included. Shorter features supplement the main contents and ensure that the most recent information available is covered. These features include research notes, book reviews, comments on previously published articles and news items about European activities in the field of agricultural economics such as meetings and conferences.

The **European Review of Agricultural Economics** is published as one volume of four issues per year (approximately 510 pages).

Subscription rates for Volume 18 (1991):

Institutions/libraries	DM 248.00 (plus postage)
Individuals (prepaid only*)	DM 120.80 (includes postage)
Single issues	DM 67.00

Prices in US\$ for subscriptions in North America only:

Institutions/libraries	US\$ 160.00 (plus postage)
Individuals (prepaid only*)	US\$ 55.20 (includes postage)

* Subscriptions for individuals are for personal use only and must be prepaid and ordered directly from the publisher. Prepayment may be made by check or by credit card: MasterCard [Access], EuroCard, Visa, and American Express [AMEX may not be used in North America]. Orders placed for institutions will be invoiced at the institutional rate. The individual rate is not available in the FRG, Switzerland, or Austria.

Institutional subscriptions and single or back issues can be ordered from your local bookseller or subscription agent, or directly from MOUTON DE GRUYTER (a division of Walter de Gruyter) at the following addresses:

For North America:
Walter de Gruyter Inc.
200 Saw Mill River Road
Hawthorne, NY 10532
USA

For all other countries:
Walter de Gruyter & Co.
Postfach 11 02 40
D-1000 Berlin 41
Federal Republic of Germany